

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: November 27, 2002, 02:06:46 ; Search time 2611 Seconds  
(without alignments)  
12104.805 Million cell updates/sec  
Title: US-09-374-967-1  
Perfect score: 1086  
Sequence: 1 atgaagccctcattctgt.....agcctgagatgcgtcatgtga 1086

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg.hum.\*
- 31: em.htg.inv.\*
- 32: em.htg.other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pln.\*
- 35: em.htg.rod.\*
- 36: em.htg.mam.\*
- 37: em.htg.vrt.\*
- 38: em.sy.\*
- 39: em.htgo.hum.\*
- 40: em.htgo.mus.\*
- 41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	692.4	63.8	1440	8	AF428297	Arabidops
2	690.8	63.6	1086	8	AY133643	Arabidops
3	690.8	63.6	1450	8	AF361812	Arabidops
4	690.8	63.6	1477	8	AF108660	Arabidops
5	690.8	63.6	1480	8	AY087698	Arabidops
6	690.8	63.6	1482	8	AY057541	Arabidops
7	690.8	63.6	1490	8	AF076484	Arabidops
8	653	60.1	142885	8	AP004127	Oryza sat
9	653	60.1	149145	8	AP003251	Oryza sat
10	651.4	60.0	142024	2	AP004705	Oryza sat
11	650.8	59.9	1478	6	AF7607	Sequence 1
12	650.8	59.9	1478	8	AF022716	Solanum t
13	638	58.7	1471	8	AB066279	Nicotiana
14	486	44.8	92624	8	AC003000	Arabidops
15	461.2	42.5	103240	8	ATT22E16	Arabidops
16	461.2	42.5	103785	8	ATF1116	Arabidops
17	451	41.5	9423	8	ATH275979	Arabidops
18	417.2	38.4	1656	3	AY071411	Drosophil
19	417.2	38.4	1715	3	AY061013	Drosophil
20	399.6	36.8	1498	9	AF135421	Homo sapi
21	399.6	36.8	1555	9	BC001141	Homo sapi
22	399.6	36.8	1555	9	BC008033	Homo sapi
23	399.6	36.8	6077	9	AB058754	Homo sapi
24	372.4	34.3	1619	8	AF234177	Pichia an
25	366	33.7	1785	8	HJ089991	Hypocrea je
26	365.8	33.7	3267	8	SCU19608	Saccharomyc
27	365.8	33.7	4446	8	SCYDL055C	S.cerevisia
28	364.2	33.5	2745	8	SCU24437	Saccharomyc
29	358	33.0	1092	6	AX250395	Sequence
30	358	33.0	15252	8	SPCC1906	Sequence
31	357.4	32.9	3220	9	AK024319	S.pombe c
32	356.2	32.8	2179	8	AB020570	Candida g
33	354.8	32.7	1275	8	D89128	Schizosacch
34	350.8	32.3	134784	8	F17123	Arabidops
35	350.8	32.3	194862	8	ATCHR173	Arabidops
36	312.4	28.8	1583	8	AF030299	Candida a
37	310.8	28.6	1089	6	AX488923	Sequence
38	306	28.2	1593	8	AF030300	Candida a
39	304.4	28.0	1665	8	AB020596	Candida a
40	282.2	26.0	144483	2	AC091701	Trypanoso
41	269.4	24.8	40519	3	LMFL2581	Leishmani
42	261.4	24.1	44686	3	AF043695	Caenorhab
43	257	23.7	2199	3	LME292039	Leishmani
44	240	22.1	267509	3	CNS07EGH	chromosom
45	215.6	19.9	45701	2	AC014674	Drosophil

ALIGNMENTS

RESULT 1  
AF428297  
LOCUS AF428297  
DEFINITION Arabidopsis thaliana At2g39770/T517.7 mRNA, complete cds.  
ACCESSION AF428297  
VERSION AF428297.1  
KEYWORDS FLI\_CDNA.  
SOURCE Arabidopsis thaliana  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi  
1 (bases 1 to 1440)  
REFERENCE Chouk.R., Chen.H., Kim.C.J., Koesema.E., Meyers.M.C., Banh.J.,  
AUTHORS Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,



RESULT 2	AY133643	1086 bp	linear	PLN 07-AUG-2002	
LOCUS	Arabidopsis thaliana At2g39770/T517.7 mRNA, complete cds.				
DEFINITION	Arabidopsis thaliana At2g39770/T517.7 mRNA, complete cds.				
ACCESSION	AY133643				
KEYWORDS	FLI_CDNA.				
SOURCE	thale cress.				
ORGANISM	Arabidopsis thaliana				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Brassicales; Brassicaceae; Arabidopsis.				
AUTHORS	1 (bases 1 to 1086) Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Banh,J., Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.				
TITLE	Arabidopsis ORF clones				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 1086) Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Banh,J., Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.				
AUTHORS	Direct Submission Submitted (17-JUL-2002) Salk Institute Genomic Analysis Laboratory (SIGnAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA				
TITLE	RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J., Sakou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.				
JOURNAL	The Salk, Stanford, PGEC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAFL cDNAs: Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Banh,J., Bowser,L., Chan,M.M., Chang,E., Dale,J.M., Deng,J.M., Goldsmith,A.D., Jones,T., Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Davis,R.W., Theologis,A., and Ecker,J.R.				
COMMENT	Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk) contributed equally to this work as PIs.				
FEATURES	Location/Qualifiers				
source	1..1086 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /chromosome="2" /clone="U14534" /note="This clone is in PUNI 51 ecotype: Columbia"				
CDS	1..1086 /note="GDP-mannose pyrophosphorylase" /codon_start=1 /product="At2g39770/T517.7" /protein_id="RAM91473.1" /db_xref="GI:22137256" /translation="MKALILVGGFGTRLRPLTLSPFKPLVDFANKPMLHQIEALKAVGVDFVLAINQPVLMNLFKDFETKLEIKITCSQETPLGTAGPLARKLDDSGEPFVLNSDVISEVPLKEMLEFHKSHGGEASIMVTKVDEPSKYGVVMEESTGRVEKF				

VEKPKLYVGNKINAGIYLLNPNSVLDKIELRPTSEIKETFFKIAAQQLYAWLPGFWM DIGQPRDYITGLRLYLSLRKSPKSLTSGPHVNVLTDEATIGSGCLIGPDVAIG PGCIVESGVRLSCTVMRGVRIKKHACISIIIGHWSVTQGWARIENMTILGEDVHVS DEIYNGGVLPKHEIKSNILKPEIYM"	BASE COUNT	299 a	220 c	285 g	282 t
ORIGIN					
Query Match	63.6%	Score	690.8	DB 8	Length 1086;
Best Local Similarity	77.3%	Pred. No.	3.8e-204;		
Matches	839;	Conservative	0;	Mismatches	247;
				Indels	0;
				Gaps	0;
QY	1	ATGAAGCCCTCATCTCTGCGGGGGTTTCGGAACCCGCTTCGGCCCTTGACTCTGAGC	60		
Db	1	ATGAAGCACTCATCTCTGTTGGAGGCTTCGGCACTCGCTTGAGACCAATTGACTCTCAGT	60		
QY	61	TTCCCGAAACCCCTCGTGGATTTTGCACCAAGCCCATGATTCGCACCAATCGAAGCT	120		
Db	61	TTCCCAAGCCCTCTGTGTGATTTTGCTAATAAACCCATGATCTTCATCAGATAGAGCT	120		
QY	121	TTGAAAGAAGTTGGGTCACAGAGGTGTTTGGCTATCAACTATCGCCCAAGAGTAATG	180		
Db	121	CTTAAGGCAGTTGGAGTTGATGAAGTGGTTTGGCCATCAATTATCAGCCAGAGGTGATG	180		
QY	181	ATTAATTTCTGAAGGACTTTGAGATAAGCTTGGCATCAACATTCATGCTCCCAAGAG	240		
Db	181	CTGAACCTCTTGAAGGACTTTGAGACCAAGCTGGAATCAAAATCACTTGTCTCAAGAG	240		
QY	241	ACTGAGCCCTTAGAAGCCGCTGGCCCTTCCTCTAGCAAGGACCAAGCTTCGGATGGA	300		
Db	241	ACCGAGCACTAGGTACCGCTGGCTCTAGCGAGAGACAAATTCCTTGATGGA	300		
QY	301	TCGCGCCAGCATCTCTTGTCTCAACAGTGTATGATAAGCAATACCCATTTGCTGAA	360		
Db	301	TCGAGAGCCCTCTTGTCTTAAACAGTGTGATTAGTAGTACCTCTTAAGAA	360		
QY	361	CTCATCAAAATTCACAAAGTGTATGGTGGAGCAACAATATGGTCACTPAAGTGGAT	420		
Db	361	ATGCTTGAGTTTCAAAATCTACGGTGGGGAAGCTCCATAATGGTAACAAAGTGGAT	420		
QY	421	GAACCATCAAAATACGGTGTGTGTTATGGAGGAGCACTGGCAGGGTGGAAAGGTTT	480		
Db	421	GAACCGTCGAAATATGAGTGGTGTATGGAAGAAAGCACTGGAAGAGTGGAAAGTTT	480		
QY	481	GTTGAGAAGCCAAAATATTGTTGGTTAAACAAGATCAATGCTGGATTACTTACTGAAC	540		
Db	481	GTGAAAGCCAAACTGTATGTAGGTACAAAGATCAACGCTGGGATTTACTTCTGAAC	540		
QY	541	CCATCTCTCTTGACCCGATTTGAGCTGAGGCCAACATCAATTTGAGAAAGAGTCTTCCT	600		
Db	541	CCATCTCTCTTGATAAGATTGAGCTAAGACCGACTTCAATCGAAAAAGAGACTTCCCT	600		
QY	601	CAATTCAGCTGATCAACAGCTCTATGCAATGCTCTCCAGGTTTTCAGGTGGATGATTT	660		
Db	601	AAGATTTCAGCAGCGCAAGGCTCTATGCTATGGTGTACCAAGGTTTTCAGGTGACATT	660		
QY	661	GGTCAGCTAGGACTACATTACTGGCTTGGCTTTTATCTAGACTCGATTAGGAAGAAA	720		
Db	661	GGCAACCCGCTGACTACATAACGGGTTTGAGACTCTACTTAGACTCCCTTAGGAAGAAA	720		
QY	721	TCAGCTGCCAAGCTAGCTACTTGAGACACATGTTTGTGGCAATGTCTGGTGCATGAGAGC	780		
Db	721	TCCTCTGCCAAATTAACCAAGTGGGCCACACATAGTTGGGAATGTTCTTGTGACGAAC	780		
QY	781	GCCAAAGATTGGGAAGTTGCTGATTTGGCTCTGATGTGGCCATTTGACCTGGGTGTT	840		
Db	781	GCTACAATTTGGGAAGGATGTTTATTTGGACACAGCGTTGCCATTTGGTCCAGGCTG	840		
QY	841	GTGGAGACCGCTGAGGCTTTCCCGCTGCACCTCTCATGCGCGCTCGCTATCAAGAG	900		
Db	841	GTTGAGTCAGAGTCAGACTCTCCCGATGCACGCTCATGCGTGGAGTCCGATCAAGAG	900		
QY	901	CATGCTTGCACTCTCAACACAGCATTTATCGGTGGGCACTCAACTGTTGGTCAATGGCAGG	960		





Db 879 TCTCTGCCAAATTAACAGTGGGCCACACATAGTTGGGAATGTTCTTGTGACAAACC 938  
Qy 781 GCCAAGATTGGAGAAGTTGTCTGATTGGTCTGATGTCGCGCATTTGGACCTGGGGTGTGT 840  
Db 939 GCTACAATTGGGAAGGATGTTGATTGGACGACGTCGTCATTTGGTCCAGGCTGCATT 998  
Qy 841 GTGAGGACGGCGTGAGGCTTTCCGCTGCACTGTCATGCGCGGGTGGCGTATCAAGAAG 900  
Db 999 GTTGAGTCAGGAGTCAGACTCTCCGATGCGCGGTGTCATGCGTGAGTCCGATCAAGAAG 1058  
Qy 901 CATGCTTGCATCTCAACAGCATTTATCGGCTGGCCTCAACTGTCGTCATTTGGTGGCCACGG 960  
Db 1059 CATCGGTGATCTCGAGCAGTATCATCGGTGGCCTCAACGGTGGTCAATGGGCCAGG 1118  
Qy 961 ATAGAGATATGACTATCTCTCGGGAGGATGTTGATGTCGTGTGATGAGTGTACAGCAAT 1020  
Db 1119 ATCAGAACATGACATCTCGGTGAGGATGTTGATGTCGCGGATGAGATCTATAGCAAT 1178  
Qy 1021 GCGCGTGTGTTTCCACATAAAGAGATCAAGTCAAGCATCTGGAAGCCGTGAGATCGTC 1080  
Db 1179 GGAGGAGTGTGTTGGCCACACAAGGAGATCAAAATCAACATCTTGAAGCCAGAGATAGTG 1238  
Qy 1081 ATGTGA 1086  
Db 1239 ATGTGA 1244  
  
RESULT 4  
AF108660  
LOCUS AF108660 1477 bp mRNA linear PLN 13-JAN-1999  
DEFINITION Arabidopsis thaliana Cytl protein (cytl) mRNA, complete cds.  
ACCESSION AF108660  
VERSION AF108660.1 GI:4151924  
KEYWORDS  
SOURCE Arabidopsis thaliana.  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
1 (bases 1 to 1477)  
Lukowitz,W. and Somerville,C.  
Unpublished  
Positional cloning of the Arabidopsis cytl gene  
2 (bases 1 to 1477)  
Lukowitz,W. and Somerville,C.  
Direct Submission  
Submitted (21-NOV-1998) Department of Plant Biology, Carnegie  
Institution of Washington, 260 Panama Street, Stanford, CA  
94305-1297, USA  
  
FEATURES  
source Location/Qualifiers  
1..1477  
/organism="Arabidopsis thaliana"  
/cultivar="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="2"  
/clone="EST T21688"  
1..1477  
/gene="cytl"  
151..1236  
/gene="cytl"  
/notes="putative mannose-1-phosphate guanylttransferase"  
/codon\_start=1  
/product="Cytl protein"  
/protein\_id="A004627.1"  
/db\_xref="GI:4151925"  
/translation="MKALILVGGFTRLRPLTILSFPPKPLVDFANKPMLHIOIEALKAV  
GVDEVLAINQPEVLMFLKDETKLEIKITCSQETPLGTAGPLAKRLDLDGSG  
EPFVLNSDVISYPLKEMLEFHKSHGSEASIMVTKVDEFSKYGVVMEESTGRVEKF  
VERPKLYGNKINAGLYLNLPSVLDKIELRPTIEKTEFFPKIAAQQGLYAMVLPFWN  
DIGQPRDITGLRLYLDLSRLKSPAKLTSGPHIVGNVLVDETATIGEGCLIGPDVAIG  
PGCIVSGVRLSCTVMRGVRIKKHACISSIIIGWHSTVGQWARIENNTILGEDVHVS  
DEIYSGGVLPKHKEIKSNILKPEIVM"

BASE COUNT 397 a 291 c 344 g 445 t  
ORIGIN  
Query Match 63.6%; Score 690.8; DB 8; Length 1477;  
Best Local Similarity 77.3%; Pred. No. 4e-204;  
Matches 839; Conservative 0; Mismatches 247; Indels 0; Gaps 0;  
Qy 1 ATGAAGCCCTCATTTCTTCGGGGGTTTCGGAAACCGCTTCGGCTTTGACTCTGAGC 60  
Db 151 ATGAAGGCACCTCATTTCTTCGGAGGCTTCGGCACTCGCTTGAGACCATTTGACTCTCAGT 210  
Qy 61 TTCCCGAAACCCCTCTGTGATTTTGCAACAAGCCCATGATTTCTGCACACAGATTCGAAGCT 120  
Db 211 TTCCCAAGACCCCTCTGTGATTTTGCTAATAAACCCCATGATCTTCATCAGATAGAGCT 270  
Qy 121 TTGAAGAAGTTGGGGTTCAGAGGTGGTTTGGCTTATCAACTATCGCCACAGAGTAAATG 180  
Db 271 CTTAAGGCAGTTGGAGTTGATGAAGTGGTTTGGCCATCAATTTATCAGCCAGAGGTGATG 330  
Qy 181 ATTAATTTCTGAAGGACTTTGAGGATAAGCTTTGGCATCACAAATATCATCTCCCAAGAG 240  
Db 331 CTGAACCTTCTGAAGGACTTTGAGACCAAGCTGGAATCAAAATCACTTCTCACAAGAG 390  
Qy 241 ACTGAGCCCTTAGGAACCGCTGGCCCTCTTGTCTTAGCAAGGACAAAGCTTGGCGATGA 300  
Db 391 ACCGAGCCACTAGGTACCGCTGGCTCTTGGCTCTTAGCGAGACAAATTTGCTTGTATGA 450  
Qy 301 TCTGCCAGCCATTTCTTGTCTCAACAGTGTATAGCGAATACCAATTTGCTGAA 360  
Db 451 TCTGGAGACCCCTCTTGTCTTAACAGTGTATAGTAGTACCTCTTTAAAGAA 510  
Qy 361 CTCATCAAAATTCACAAGTGTCTATGTTGGTGGAGCAACAATTTATGGTCACTAAGGTGGAT 420  
Db 511 ATGCTTGAGTTTCACAAATCTCACGCTGGGAGCCCTCCATAATGTTAACAAGGTGGAT 570  
Qy 421 GAACCATCAAAATACGGTGTGTGGTTATGGAGGAGGCAACTGGCAGGTGGGAAAGATTT 480  
Db 571 GAACCGCTGAAATATGAGTGGTGTGTATGGAAGAAAGCACTGGAAGAGTGGGAAAGATTT 630  
Qy 481 GTTGAGAACCCAAAATATTTGTGGTCAACAGATCAATGCTGGGATTTACTTACTTGAAC 540  
Db 631 GTGGAAGCCAAAATGTATGTAGTAAACAGATCAACGCTGGGATTTATCTTCTGAAAC 690  
Qy 541 CCATCTGTCTTACCGCATTTGAGCTGAGGCCAACATCAATTCAGAAAGAGGTCTTCCCT 600  
Db 691 CCATCTGTCTTGATAAGATTGAGCTAAGCCGACTTCAATCGAAAGAGAGACTTTCCCT 750  
Qy 601 CAAATTCAGCTGATCAACAGCTCTATGCAATGGTCCCTTCCAGGTTTGGATGGATGTT 660  
Db 751 AAGATTGCAGCAGCGCAAGGGCTCTATGCTATGGTGTCTACCAAGGTTTGGATGGACATT 810  
Qy 661 GGTACGCCCTTAGGACTACATTTACTGGCTTGGCTTATCTAGACTCGATTAGGAAGAAA 720  
Db 811 GGGCAACCCGCTGACTACATAAGCGGTTTGAGACTCTACTTACACTCCCTTAGGAAGAAA 870  
Qy 721 TCAGCTGCCAAGCTAGCTACTGGAGCACATGTTGTTGGCAATGCTGTGGTGCATGAGAGC 780  
Db 871 TCTCTGCCCAATTAACAGTGGGCCACACATAGTTGGGAATGTTCTTGTGAGGAACC 930  
Qy 781 GCCAAGATTGGAGAAGTTGTCTGATTGGTTCCTGATGTGCCCATTTGGACCTGGGTGTGTT 840  
Db 931 GCTACAATTTGGGAAGGATGTTTGAATTGGACCAGACGTTGCCATTTGTCACAGCTGCATT 990  
Qy 841 GTGGAGGACGGGTGAGGCTTTCCCGCTGCAGCTGTGATGCGCGGGCTGCGCTATCAAGAAG 900  
Db 991 GTTGAGTCAGGAGTCAGACTCTCCCGATGACGCTCATGCTGGAGTCCGATCAAGAAG 1050  
Qy 901 CATGCTTGCATCTCAACAGCATTTATCGGCTGCACATCAACTTGTGTCATGGCACGG 960  
Db 1051 CATCGGTGTATCTCGAGCAGTATCATCGGCTGGCAGTCAACAGGTTGGTCAATGGGCCAG 1110  
Qy 961 ATAGAGAAATATGACTATCTCTGGGGAGGATGTTTCATGTGTGATGAGGTGTACAGCAAT 1020



Db	1115	ATCAGAACATGACGATCCTCGGTGAGGATGTTTCATGTGAGCGGATGAGATCTATAGCAAT	1174
QY	1021	GGCGGTGTTTCCACATPAAAGAGATCAAGTCAAGCATTTCTGAAGCCTGAGATCGTC	1080
Db	1175	GGAGGAGTGTGTTTGGCCACACAGGAGATCAAAATCAACATCTTGAAGCCAGAGATAGTG	1234
QY	1081	ATGTCGA	1086
Db	1235	ATGTGA	1240
RESULT	6		
AY057541			
LOCUS		1482 bp mRNA linear	PLN 07-OCT-2001
DEFINITION		Arabidopsis thaliana At2g39770/T5I7.7 mRNA, complete cds.	
ACCESSION		AY057541	
VERSION		AY057541.1	GI:15982867
KEYWORDS		FLI CDNA.	
SOURCE		Arabidopsis thaliana.	
ORGANISM		Arabidopsis thaliana	
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			

1..158

159..1244

/note="GDP-mannose pyrophosphorylase"

/codon\_start=1

/product="At2g39770/T5I7.7"

/protein\_id="AAL09781.1"

/db\_xref="GI:15982868"

/translation="MKALILVGGFGTRLRPLPLSPKPLVDFANKPMILHQIETALKAV

GYDEVVLAINIQPEVLAFLKDFETKLEIKITCSQETEPPLGTAGPLAARDLLDSSG

EPFVLNSDVISEYPLKEMLEPHKSHGGEASIMVTQKDEPESKYGVVYMLARESTGRVEKF

VEKPKLYVGNKINAGIYLLNPVLDKIELRPTSIEKETFPKIAAAGLYAMVLPGFWM

DIGQPRDYITGLRLYLSLRKSPAKLTSRPHIVGNLVDETATGEGCLIPDVAIG

PECIVESGVRUSCTGVMRGVRIKKHACISSIIHSHSTVQWARIENMTLILGEDVHVS

DEIYSGNGVWLPKRIKSNILKPEIVM"

1245..1482

398 a 293 c 345 g 446 t

3'UTR

BASE COUNT

ORIGIN

Query Match 63.6%; Score 690.8; DB 8; Length 1482;

Best Local Similarity 77.3%; Pred. No. 4e-204;

Matches 839; Conservative 0; Mismatches 247; Indels 0; Gaps 0;

QY 1 ATGAAGGCCCTCATTTCTTCGGGGGTTTCGGAACCCGCTTCGGCCTTTGACTCTCAGC 60

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 159 ATGAAGGCACTCATTTCTTTGGAGGCTTCGGCACTCGCTTGAGACCAATTTGACTCTCAGT 218

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 61 TTCCCCAAAACCCCTCTCGTATTTTGCACAAACAGCCCATGATTTCTGCACACGATCGAAGCT 120

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 219 TTCCCCAAAACCCCTCTGATTTTGTATTAATAAACCCATGATCTCTCATCAGATAGAGGCT 278

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 121 TTGAAAGAAGTTGGGTCACAGAGTGGTTTGGCTATCAACTATCGCCACGAGGTAATG 180

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 279 CTTAAGCGAGTTGGAGTTGATGAAGTGGTTTGGCCATCAATTTATCAGCCAGAGGTATG 338

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 181 ATTAATTTCTGAAGACATTTGAGATAGCTTGGCATCAACAATTACATGCTCCCAAGAG 240

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 339 CTGAACCTTTCTGAAGACATTTGAGACCAAGCTGGAATCAAAATCACTTCTCACAAGAG 398

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 241 ACTGAGCCCTTAGGAACCGCTGCGCTCTAGCAGGACAGCTTGGCGATGGA 300

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 399 ACCGAGCCACTAGTACCGCTGCTCTGCGCTCTAGCAGAGACAAATTTGCTTGATGGA 458

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 301 TCTGGCAGCCCATTTCTTCTCAACAGTGTATGTCATAAGCGAATACCCATTTGCTGAA 360

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 459 TCTGGAGACCCCTCTCTTCTTACAGTGTATGTATGAGTACAGTACCTCTTAAGAA 518

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 361 CTCATCAAAATTCACAAGTGTATGTTGAGGCAACAATATGTTACTAAGGTGGAT 420

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 519 ATGCTTGAGTTTCACAAATCTCACGTTGGGGAAGCCTCCATAATGGTAACAAAGGTGAT 578

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 421 GAACCATCAAAATACCGTGTGTTGTTATGGAGGAGCAACTGGCAGGTGGAAAGGTTT 480

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 579 GAACCGTCAAAATATGAGTGGTTGTTATGGAAAGAACCACTGGAAGAGTGGAGAAAGTTT 638

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 481 GTTGAAGACCCAAAATATTTGTGGTAAACAAGTCAATGCTGGGATTTTACTTACTGAAC 540

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 639 GTGGAAGACCCAAAATCTGTATGAGTACACAAGATCAACCTGGGATTTATCTTCTGAAC 698

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 541 CCATCTGTCTTTGACCGCATTTGAGCTGAGGCCAACATCAATTCAGAAAGAGGCTTCCCT 600

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 699 CCATCTGTCTTTGATAAAGTTGAGCTAAGACCGACCTCAATCGAAAGAGAGACTTCCCT 758

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 601 CAATTTGACGTGATCAACAGCTCTATGCAATGGTCTTCCAGGTTTTTGGATGGATGTT 660

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 759 AAGATTGACAGCGCAAGGCTCTATGCTATGGTCTACAGGTTTTTGGATGGACATT 818

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 661 GGTACGCTAGGACTACATTTACTGGCTTGGCTTTTATCTAGACTCGATTAGGAGAAA 720

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 819 GGGCAACCCCGACTACATACGCGGTTTGAGACTCTACTTAGACTCCCTTTAGGAGAAA 878

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 721 TCAGCTGCCAAGCTAGCTACTTGGAGCACATGTTGTTGGCAATGTGCTGTCATGAGAGC 780

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 879 TCTCTGCGCAATTAACCAAGTGGCCACACATAGTTGGGATGTTCTTGTGACGAACC 938

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||



QY	721	TCAGCTGCCAAGCTAGCTACTGAGCACATGTTGTTGGCAATGCTGCTGGTCATGAGACC	780
Db	873	TCTCCTGCGCAATTAACACAGTGGGCCACACATAGTTGGGAATGTTCTTTGTTGACGAAACC	932
QY	781	GCCAAAGATTGGAGAGGTTGCTGATTGCTGCTGATGTCGCCATTGGACCTGGGTGTGT	840
Db	933	GCTACAATTGGGAAGGATGTTGATTGGACAGACAGTTGCCATTGGTCCAGGCTGCATT	992
QY	841	GTGAGGACGCGGTGAGCGTTTCCGCGTGCACGTGTCATCGCGCGGTGCGGTATCAAGAAG	900
Db	993	GTTGAGTCAGGAGTCAGACTCTCCGATGCACGGTTCATGCGTGGAGTCCGCATCAAGAAG	1052
QY	901	CATGCTGCATCTCAACAGACATATATCGCTGGCGCTCAACTGTTGGTCAATGGCGACGG	960
Db	1053	CATCGTGTATCTCAGCAGATATCATCGGGTGGCACTCAACGGTTGGTCAATGGGCCAGG	1112
QY	961	ATAGAGATATGACTATCTCTGGGGAGGATGTTCTATGTTGTGTGAGGAGTGTACAGCAAT	1020
Db	1113	ATCAGAAATCATGACATCTCTCGGTGAGGATGTTCTATGTGAGCGATGAGATCTATAGCAAT	1172
QY	1021	GGCGGTGTTGTTCTCCACATAAAGAGATCAAGTCAAGCATCTCTCAAGCCTGAGATCGTC	1080
Db	1173	GGAGGATGTTTTCACACACAGGAGATCAATCAACATCTTCAAGCCAGATAGTG	1232
QY	1081	ATGTGA	1086
Db	1233	ATGTGA	1238
RESULT 8			
AP004127			
LOCUS			
DEFINITION			
Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,			
clone:P0005H10.			
ACCESSION			
AP004127			
VERSION			
AP004127.1 GI:15408872			
KEYWORDS			
SOURCE			
Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,			
clone:P0005H10.			
ORGANISM			
Oryza sativa (japonica cultivar-group)			
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
Ehrhartoideae; Oryzeae; Oryza.			
REFERENCE			
1			
Sasaki, T., Matsumoto, T. and Yamamoto, K.			
Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC			
clone:P0005H10			
JOURNAL			
Published Only in Database (2001)			
REFERENCE			
2 (bases 1 to 142885)			
Sasaki, T., Matsumoto, T. and Yamamoto, K.			
Direct Submission			
TITLE			
Submitted (31-AUG-2001) Takuji Sasaki, National Institute of			
Agrobiological Sciences, Rice Genome Research Program; Kannondai			
2-1-2, Tsukuba, Ibaraki 305-8602, Japan			
(E-mail:tsasakienias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,			
Tel:81-298-38-7441, Fax:81-298-38-7468)			
Genes were predicted from the integrated results of the following:			
GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor			
(October 1998 version). The genomic sequence was searched against			
NCBI Nonredundant Protein database, nr			
(ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at			
RGP. Protein homologies of the coding regions were searched against			
NCBI Nonredundant Protein databases using BLASTN 2.0. ESTs represent			
the identified cDNA sequences using BLASTN 2.0 with the			
corresponding DDBJ accession no. and RGP clone ID.			
A gene with identity or significant homology to a protein is			
classified based on the protein name to indicate the homology level			
such as same name, 'putative-' and '-like protein'. A gene without			
significant homology to any protein but with EST homology (covering			
almost the entire length of partial sequence) is classified as an			
'unknown' protein. A gene predicted with a gene prediction program			
is classified as a 'hypothetical' protein.			
The orientation of the sequence is from SP6 to T7 of the PAC clone.			

Detailed information on overlap and assembly quality together with annotation of this entry is available at http://rgp.dna.affrc.go.jp/GenomeSeq.html.			
Location/Qualifiers			
1. .142885			
/organism="Oryza sativa (japonica cultivar-group)"			
/cultivar="Nipponbare"			
/db_xref="taxon:39947"			
/chromosome="1"			
/clone="P0005H10"			
complement(join(3920. .3951,4840. .4981))			
/gene="P0005H10.1"			
complement(join(3920. .3951,4840. .4981))			
/gene="P0005H10.1"			
/note="hypothetical protein"			
/codon_start=1			
/protein_id="BAB64264.1"			
/db_xref="GI:15408873"			
/translation="MRGGAAAGYAEAGSAGVVEEPPRRVVRKRREEVEAACTMA			
IGESWRTLIQYIF"			
join(6179. .6332,7620. .7781,8118. .8296,8558. .8683,			
8722. .8853,9433. .9540,9645. .9863,9939. .10119,10671. .10792)			
/gene="P0005H10.2"			
join(6179. .6332,7620. .7781,8118. .8296,8558. .8683,			
8722. .8853,9433. .9540,9645. .9863,9939. .10119,10671. .10792)			
/gene="P0005H10.2"			
/note="contains EST C26260(C11984)"			
similar to Arabidopsis thaliana chromosome 2, T517.4			
unknown protein"			
/codon_start=1			
/protein_id="BAB64265.1"			
/db_xref="GI:15408874"			
/translation="MSINFPVNVVLEKCTEDILSLIKPVEGRNKRKIYAIQELADTI			
XSAGALRGASVYFGFSVQLYAKSGDLVSVELFNALNLPISKRRKQDLTLEVRRL			
QKRGIAHMEFIPNARVPVLYVNOYISDCISISNYPGRISKIFYYWINTLDDRF			
DMVLTCEPAILLPLKEIYEGNIMEDISKISLWEYGLPMDSDYLGFRAYNEKHL			
EVCSTNIERFRQNMQRNQSILSHLASFHKFRIDALSDKIVSTYGRLEIRQDN			
PWWMKSYLFEVDPFEKPDNAARAVGSEFODIVNAFNASNKFVSDAHLTDRL			
LSLLCTPVGSKLGRASRYNTLPAEGSGSGSGPHGYTNRQTAVYONNHPO			
AYNTQRQTVVHQNNQVYAGRGRTLPARGSKNLQNGSTWRSTJUANLTTEHRNM			
HVTRRRNAL"			
complement(join(13608. .13812,14087. .14164,14524. .14646,			
15045. .15302,15753. .15912,16541. .16651,16998. .17268,			
17353. .17502,18222. .18929))			
/gene="P0005H10.3"			
complement(join(13608. .13812,14087. .14164,14524. .14646,			
15045. .15302,15753. .15912,16541. .16651,16998. .17268,			
17353. .17502,18222. .18929))			
/gene="P0005H10.3"			
/codon_start=1			
/product="ankyrin-like protein"			
/protein_id="BAB64266.1"			
/db_xref="GI:15408875"			
/translation="MKPLGAGHADLLCGPRVLRRAALAFACAAFLVGVHWSSPRL			
LFFSTASSSSPLSLSTGSPSVAVSPNANLSDPSLIPTPAASTPASPANASPPS			
LPPPPPLPPPPPPARLGIYGEGRMRDDFVVGANDTDLAANTDEALPQPTDAGP			
AVGSRVIRGRFVPCESMREYIPCLDNEEIRRLPSTEDGERFERHCPRDKGLSLV			
PAPKYKAPIPWPSRDEWFNSVPHRLVDDKGQNWISKADKFRFPFGGQTIFHG			
ANQYLDQISQWPDIAFGSHTRVLDVGGVAGFYGLLSRDVLTSLAPKDVYHENQI			
QFALRGVPMAMAAAFATRLRLYPSQAFDLIHCSRIRINWTHDDGILLVLRNMLRAG			
YFAAQAQPVYKHEEAQAEKEMEDFTARLCWELVKEGYIAMWRKPLANSVCYMRDP			
GVKPALCPDDNDPVDVYVNLKACISRLPENGSDGLTPFPWPAPRLMEPPKRLGVMDA			
VNSKRELKAEKTFKWDIDIVGIVRFVKKRKLNRVLDMRAGFGGFAAALINRKLDCW			
HNNVVPVSEPTNLPVIYDRGLGVADHCEPFTPTPTDYDLLHAFSLFKEQKRCNIS			
STILLEMDRILRPGRAYIRDLQVQVQVQVKEITAMGWSIRMRDTASGVSASRKLMD			
KPMV"			
join(22002. .22041,22106. .22284)			
/gene="P0005H10.4"			
join(22002. .22041,22106. .22284)			
/gene="P0005H10.4"			
/note="hypothetical protein"			
/codon_start=1			
/protein_id="BAB64267.1"			





QY	938	CAACTGTTGGTCAATGGCAGGATAGAGAATATGACTATATCTCTGGGGGAGGATGTTTCATG	997
Db	42607	CAACTGTTGGAGATGGGCAGGATAGAAATCTGACTATCTCTGGGAGGACGTACATG	42666
QY	998	TGTTGTGATAGGTGTACAGCAATGCGGTGTGTTCTCCACATAAAGAGATCAAGTCAA	1057
Db	42667	TAGTGTGATGAGTCTATACCAACGCGGTGTGTTCTCCGCACAAAGAGATCAAGTCAA	42726
QY	1058	GCATTCTGAAGCCTGAGATCGTCATGTGA	1086
Db	42727	GCATCTGAAGCCTGAGATTGTCATGTGA	42755
RESULT 9			
AP003251			
LOCUS		149145 bp	DNA linear
DEFINITION		Oryza sativa (japonica cultivar-group)	genomic DNA, chromosome 1,
ACCESSION		PAC clone:P0446B05.	
VERSION		AP003251	
KEYWORDS		AP003251.3	GI:20160607
SOURCE			
ORGANISM		Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,	
		clone:P0446B05	
		Oryza sativa (japonica cultivar-group)	
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
		Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;	
		Ehrhartoideae; Oryzeae; Oryza.	
REFERENCE		1	
AUTHORS		Sasaki, T., Matsumoto, T. and Yamamoto, K.	
TITLE		Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC	
		clone:P0446B05	
JOURNAL		Published Only in Database (2001)	
REFERENCE		2 (bases 1 to 149145)	
AUTHORS		Sasaki, T., Matsumoto, T. and Yamamoto, K.	
TITLE		Direct Submission	
JOURNAL		Submitted (19-FEB-2001) Takuji Sasaki, National Institute of	
		Agrobiological Sciences, Rice Genome Research Program; Kannondai	
		2-1-2, Tsukuba, Ibaraki 305-8602, Japan	
		(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,	
		Tel:81-298-38-7441, Fax:81-298-38-7468)	
COMMENT		On Apr 16, 2002 this sequence version replaced gi:17933036.	
		Genes were predicted from the integrated results of the following:	
		GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor	
		(October 1998 version). The genomic sequence was searched against	
		NCBI Nonredundant Protein database, nr	
		(ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at	
		RGF. Protein homologies of the coding regions were searched against	
		NCBI Nonredundant Protein database with BLASTP2.0. ESTs represent	
		the identified cDNA sequences using BLASTN 2.0 with the	
		corresponding DBJ accession no. and RGP clone ID.	
		A gene with identity or significant homology to a protein is	
		classified based on the protein name to indicate the homology level	
		such as same name, 'putative-' and '-like protein'. A gene without	
		significant homology to any protein but with EST homology (covering	
		almost the entire length of partial sequence) is classified as an	
		'unknown' protein. A gene predicted with a gene prediction program	
		is classified as a 'hypothetical' protein.	
		The orientation of the sequence is from SP6 to T7 of the PAC clone.	
		This orientation of P0446B05 clone has an overlap with P0406G08	
		(DBJ:AP003240) clone at the position 1 to 17,096 of 5' end and	
		with P005UH10 (DBJ:AP004127) at the position 98,039 to 149,145 of	
		3' end. Detailed information on overlap and assembly quality	
		together with annotation of this entry is available at	
		http://rgp.dna.affrc.go.jp/GenomeSeq.html.	
FEATURES		Location/Qualifiers	
source		1. .149145	
		/organism="Oryza sativa (japonica cultivar-group)"	
		/cultivar="Nipponbare"	
		/db_xref="taxon:39947"	
		/chromosome="1"	
		/clone="P0446B05"	
gene		complement(join(1108. .1218,2391. .2486,2568. .2681,	
		2756. .2833,2914. .3015,4386. .4661))	

		/gene="P0446B05.1"	
		complement(join(1108. .1218,2391. .2486,2568. .2681,	
		2756. .2833,2914. .3015,4386. .4661))	
		/gene="P0446B05.1"	
		/note="contains ESTs D23149(C2313),C26277(C12019)	
		similar to Arabidopsis thaliana chromosome 3,At3g09320	
		unknown protein"	
		/codon_start=1	
		/protein_id="BAB89554.1"	
		/db_xref="GI:20160608"	
		/translation="MAGRPGYVTPILSVLAAIGVYVYTAFLAIPAWLGLATAAGVA	
		NAVFTALAAACVATYAVAVSRDPGRVPAPFLDVEDAESPIHEIKRKHCHCWINNC	
		VGHENKIFLVEVLYAVASLSVLVIGGAVHSLPNQOLGSDSRISLVSNQIIC	
		GVFLCPALALSLILGWHVYLIFHNKTTIEYHEGVRAMWLAERKAGNLYHHPYDLGVYE	
		NIIVSVLGNALCWLCPISRNITNGIRFTSYDIPLSTPPI"	
		complement(11360. .12688)	
		/gene="P0446B05.2"	
		complement(11360. .12688)	
		/gene="P0446B05.2"	
		/note="contains EST AU176487(E21047)	
		similar to Arabidopsis thaliana chromosome 5, At5g02190	
		unknown protein"	
		/codon_start=1	
		/protein_id="BAB89555.1"	
		/db_xref="GI:20160609"	
		/translation="MPPPALPVCLLLLLLLLLLAVPRPAAAAAATRPLLFLRAR	
		QVPAGALPRPASKLRPHNVSLTVSLAVGTPPQNVTWLDTSGLSWLLCAPGGGGG	
		GGRSALSRPRASLTFAVPCDSACRSRDLPPSPACDGASKOCRSVLSQAGSSGG	
		ALATEVFTVQGPPPLRAAFGCMTAEDTSPDGVATAGLGMNRCALSVLSQASTRES	
		YICISRDDAGVLLGHSDLPFLPLNTPLQPAAMPLPYFDRVAYSVOLLLGIRVGGKPL	
		PIPASVLAFDHGTAGTMDVSGTQFTLLGDAYSAKAEFSRQTKPWLPAINDPNFAF	
		QEAFTCFRPQGRAPPARLPVLLFNQAQMTVAGDRLLYKVPGERGGDGVWMLTF	
		GNADWVPIATYVIGHHNMNVVEYDLERGRVGLAPIRCDVASERLGLML"	
		join(13981. .14067,14601. .14684,14736. .14909)	
		/gene="P0446B05.3"	
		join(13981. .14067,14601. .14684,14736. .14909)	
		/gene="P0446B05.3"	
		/note="hypothetical protein"	
		/codon_start=1	
		/protein_id="BAB89556.1"	
		/db_xref="GI:20160610"	
		/translation="MGEKCIAMGNAGSQARHRLPDVETSKDDIIYQDYTTTGS	
		LVDLQKSHCRDRVPPMPYSPGLHDALFGLDITILODEDEHNSRMRSNMRHNDRR	
		LTALIIADVTSN"	
		complement(17162. .17869)	
		/gene="P0446B05.4"	
		complement(17162. .17869)	
		/gene="P0446B05.4"	
		/note="hypothetical protein"	
		/codon_start=1	
		/protein_id="BAB89557.1"	
		/db_xref="GI:20160611"	
		/translation="MATFSTAGGIMDASPIWCFMCSRLRHPDGLSTCPTPRPAAALE	
		EIVEVMDAGEFLQACALRRAPVAAVSSSTQQLTPTVTRDAGRTCAVCLDLDEPGSA	
		VVTPCGDHAYHPQCIAPLWLEHDTCLCRRESGLQVVEVQDGMVLSLSPDLVLCLE	
		MMPGSRSEVLRGRVAGRIFAVRVVDGTGKLVKRGVLRUGSACHRFAAAAGNLLSLR	
		YRDCVIPNDLLIGVOC"	
		join(22593. .23216,24099. .25639,26226. .26367,26499. .26606,	
		26691. .26903,26990. .27100,27933. .28127,28751. .28925,	
		29066. .29103)	
		/gene="P0446B05.5"	
		join(22593. .23216,24099. .25639,26226. .26367,26499. .26606,	
		26691. .26903,26990. .27100,27933. .28127,28751. .28925,	
		29066. .29103)	
		/gene="P0446B05.5"	
		/codon_start=1	
		/product="putative pumilio/Mpt5 family RNA-binding	
		protein"	
		/protein_id="BAB89558.1"	
		/db_xref="GI:20160612"	
		/translation="MGGGGGGGDEGELELDALLGTGGAGARRRGDAGERERELNM	
		FRSGSAPPTIEGSLNIGLLRGGAATVAAITPDATETLNGHGLLSEDLRADPAY	
		LSYYYSHGNLNPRLPPVLKSDWRSTQRLKAGVVGIGDGRKVSPEETGHEPTVGRP	





DEFINITION Oryza sativa (japonica cultivar-group) chromosome 8 clone P0682A06,  
\*\*\* SEQUENCING IN PROGRESS \*\*\*, in ordered pieces.

ACCESSION AP004705

VERSION AP004705.1 GI:18447964

KEYWORDS HTG; HTGS\_PHASE2.

SOURCE Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,  
clone:P0682A06

ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE 1

Sasaki, T., Matsumoto, T. and Yamamoto, K.

Oryza sativa nipponbare(GA3) genomic DNA, chromosome 8, PAC

clone:P0682A06

Published Only in Database (2002)

2 (bases 1 to 142024)

Sasaki, T., Matsumoto, T. and Yamamoto, K.

Direct Submission

Submitted (30-JAN-2002) Takuji Sasaki, National Institute of

Agrobiological Sciences, Rice Genome Research Program; Kannondai

2-1-2, Tsukuba, Ibaraki 305-8602, Japan

(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,

Tel:81-298-38-7441, Fax:81-298-38-7468)

NOTE: It currently consists of 1 contigs. Gaps between the contigs  
are represented as runs of N. The order of the pieces is believed  
to be correct as given, however the sizes of the gaps between them  
are based on estimates that have provided by the submitter. This  
sequence will be replaced by the finished sequence as soon as it is  
available and the accession number will be preserved.

\* NOTE: This is a 'working draft' sequence.

\* This sequence will be replaced

\* by the finished sequence as soon as it is available and

\* the accession number will be preserved.

FEATURES

source

1..142024

/organism="Oryza sativa (japonica cultivar-group)"

/cultivar="Nipponbare"

/db\_xref="taxon:39947"

/chromosome="8"

/clone="P0682A06"

BASE COUNT 39960 a 30290 c 31009 g 40765 t

ORIGIN

Query Match 60.0%; Score 651.4; DB 2; Length 142024;  
Best Local Similarity 81.2%; Pred. No. 1.7e-191;  
Matches 803; Conservative 0; Mismatches 111; Indels 75; Gaps 1;

Qy 173 AGGTAATGATTAATTTCTTGAGGACTTTGAGGATAAGCTTGGCATCACAAATCACGCTGT 64921

Db 64862 AGGTAATGCTCAATTTCTGAAGGACTTTGAGGATAAGCTTGGCATCACAAATCACGCTGT 64921

Qy 233 CCCAAGACACTGAGCCCTTAGGAACCGCTGGCCCTCTGCTAGCAAGGGACAGCTTG 292

Db 64922 CCCAAGACACTGAGCCCTTAGGAACCGCTGGCCCTCTGCTAGCAAGGGACAGCTTG 64981

Qy 293 CGGATGATCTGGCCAGCCATTTCTTTGCTCCCAACAGTATGCTCATAGCGGAATACCCAT 352

Db 64982 TGGATGATCTGGTGAGCCATCTTTTGCTCCCAACAGTATGCTCATAGCGGAATACCCAT 65041

Qy 353 TTGCTGAACATCATCAAAATTTTACAAAGTGTGATGGTGGGAGGACCAAAATTTATGGTCACPA 412

Db 65042 TTGCTGAGCTCATAAAATTTACAAAGACCATGTTGGTGGGAGGACCAAAATTTATGGTCACCA 65101

Qy 413 ----- 412

Db 65102 AGGTAGATCCTTTCTTTGACCACTTTCTTTGGATTATAAACTCGTTGATTGTGTATACTGA 65161

Qy 413 -----AGGTGATGATCAACCATCAAAATACGGTGTGTTGTTATGGAGGAGG 457

Db 65162 CTATTATTTCAATGTAGTGGACGACCATCAAAATACGGTGTGTTGTTATGGAGGAGG 65221

Qy 458 CAACTGGCAGGTTGGAAAGGTTTGTGAGAACCCAAAAATATTTGTGGGTAAACAAGATCA 517

Db 65222 CCACTGGAATGGTGGAAAAATTTGTTGAGAAACCAAAAAATATTTGTAGGCAACAGATCA 65281

Qy 518 ATGCTGGGATTTACTTACTGAACCCATCTGCTCTTGACCGCATTTAGCTGAGGCGCAACAT 577

Db 65282 ATCGAGGAATTTACTTGTGTAACCCGCTCTGCTCTGGACCGCATCGAGCTAAGCCACTT 65341

Qy 578 CAATGAGAAAGAGGCTTCCCTCAAAATGCGAGCTGATCAACAGCTCTATGCAATGGTCC 637

Db 65342 CAATGAGAAAGAGGCTTCCCTCGAATCTCAGCTGATGCAAAAGCTTTTGTCTGATGGTCC 65401

Qy 638 TTCCAGGTTTTTGATGCGATGTTGGTTCAGCCTAGGGACTACATTTACTTGGCTTGGCTTT 697

Db 65402 TTCCAGGTTTTTGATGCGATGTTGGTTCAGCCTAGGGACTACATTTACTTGGCTTGGCTTT 65461

Qy 698 ATCTAGACTCGATTAGGAAGAAATACGCTGCCAAGCTAGCTTACTTGGAGCACATTTGTTG 757

Db 65462 ATCTGATTCACATTAGGAAGAGATCAGCCAAACAGGTTAGCCACTGGAGCACACATTTGTTG 65521

Qy 758 GCAATGTGCTGGTGCATGAGAGCGCCCAAGATTGGAGAAGGTTGTCTGATTGGTCTCTGATG 817

Db 65522 GGAATGTGCTAGTTCACGAGAGTGCAGAAAGATTGGCAAGGCTGTCTGATTGGTCTCTGATG 65581

Qy 818 TCGCCATTGGACCTGGGTGTTGTTGAGGACGCGCTGAGGCTTTCCCGCTGCACTGTCA 877

Db 65582 TTGCTATTGGTCTCTGGATGCTGCTGGAGGATGGTGTGAGGCTTTCCCGTTTGCACGCTGA 65641

Qy 878 TGGCGGCGTGGCTGATCAAGAAGCATCTTGCATCTCAACAGCATTTATCGCTGGCGGCACT 937

Db 65642 TGGTGGCGTGGCTGATTAAGAAGCATCTTGCATATCAACAGCATTTATTTGGATGGCACT 65701

Qy 938 CAACTGTTGGTCAATGGGACGATAGAGAAATATGACTATCTCTGGGGAGGATGTTTCATG 997

Db 65702 CAACTGTTGACAAATGGGACGATAGAGAAATATGACTATCTCTAGGAGGACGTACATG 65761

Qy 998 TGTGTGATGAGGTGTACAGCAATGGCGGTGTTGTTCTCCCAATPAAAGAGATCAAGTCAA 1057

Db 65762 TAGGTGATGAGGTCTATACCAACGGCGGTGTTATTTCTCCGACAAAGAGATCAATCAA 65821

Qy 1058 GCATTCTGAAGCCTGAGATGCTCATGTA 1086

Db 65822 GCATCCTGAAACCTGAGATTGTCTATGTA 65850

RESULT 11

A97607

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1..1478

/organism="Solanum tuberosum"

/db\_xref="taxon:4113"

/tissue\_type="BLATTGEWEBE"

/clone\_lib="CDNA-BIBLIOTHEK AUS BLATTGEWEBE"

/clone\_lib="(B) CLON(E): PFMFASE"

/notes="unnamed protein product"

/codon\_start=1

/protein\_id="CAB70521.1"

CDS

```
/db_xref="GI:6780909"
/translation="MKALILVGGFOTRLRPLPLSVKPKPLVEFANKPMILHQIEALKAV
GTYEVLAINYOEVLNFKFEASLGKITCSQETPLGTAGPLALARDKLIDDSG
EPFVLNSDVISSEYFKEMIOFPHKSHGFEASLMVTKVDEPSKYGVVYMEESTGOVERF
VPKPLVQNKINAGPVLNPSVLDRIOLRPHSIEKEVPPKIAAEKKLYAMVLDPFWM
DZGQPRDITGLRLDLSDSKNISSFKLASGPHIVGNYIVDESARKIGEGCLIPDVAIG
SGCVIESGVRUSRCTVMRGVRIKKHAKICSGSIIGHWSTVGOWARVENMTILGEDVHVC
DEIYSGGVLPKHKEIKSILKPEIVM"
BASE COUNT      395 a      279 c      334 g      470 t
ORIGIN

Query Match      59.9%; Score 650.8; DB 6; Length 1478;
Best Local Similarity 75.0%; Pred. No. 1.3e-191;
Matches 814; Conservative 0; Mismatches 272; Indels 0; Gaps 0;

QY 1 ATGAGGCCCTCATCTTGTGCGGGGTTTCGGAACCGCTTCGCGCTTTGACTCTGAGC 60
Db 136 ATGAAGCCATATCCTGTTGGAGGGTTGCGGTACTGCGGTACAGGCCACACCCCTCAGT 195

QY 61 TTCCCGAAACCCCTCGTGGATTTTGCACAAAGCCCATGATTTGCACACGATCGAAGCT 120
Db 196 GTCCCAAGCCACTCGTCGAAATTTGCTAACAAACCAATGATTTTGCATCAGATTGAGGCT 255

QY 121 TTGAAAGAAAGTGGGGTCCACAGAGTGGTTTGGCTATCAACTATCCACCAAGCTGAATG 180
Db 256 CTCAGGCTGTTGGAGTAAACCAAGTTGTACTGGCTATTAACTACCACCAAGCTGAGGTGATG 315

QY 181 ATTAATTTCTTGAAGACATTTGAGGATAAGCTTTGGCATCACAAATTACATGCTCCCAAGAG 240
Db 316 CTGAAGCTTTTCAAGAAATTTGAGGAAGCCCTTGGAAATCAAGATCAAGCTGTTCTCAAGAA 375

QY 241 ACTGAGCCCTTAGGAACCGCTGCGCTCTTGTCTAGCAAGGACCAAGCTTGCAGATGGA 300
Db 376 ACTGAACCACTTGGCACTGCTGCTGCTGCTGCTTGGCTAGAGTAAGCTAGTAGATGAC 435

QY 301 TCTGCCACCAATTCCTTCTCAACAGTGATGTCATGAAGCAATACCAATTCCTGCTGAA 360
Db 436 TCTGGTGAACCATTTTCTTCTTAAAGTGATGTTATCAGTGAATATCCATTCAGGAG 495

QY 361 CTCATCAAAATTCACAAATGTCATGTTGAGGAGGCAACATGATGTCACATAGGTGGAT 420
Db 496 ATGATTCAAATTCACAAATCCCAATGCTGTTGAGGCTTCTTGTGATGTCACCAAGGTGGAT 555

QY 421 GAACCATCAAAATACGGTGTGTGGTTTATGGAGGAGGCAACTGGCAGGTGGAAAGGTTT 480
Db 556 GAGCCTTCTAAATATGCTGTTGTGTCATGGAAGCAATCCACTGCGGCAAGTAGACATTT 615

QY 481 GTTGAAGCCCAAAAATATTTCTGCGCTTAAACATCAATCTCTGCGCATTTACTTACTCAAC 540
Db 616 GTGGAGAAGCCAAAGTTATTTGTGGCAACAAGATCAATGCTGATTTTACCTGCTGAAC 675

QY 541 CCATCTGCTCTGACCGCATTTAGCTGAGGCGCAACATCAATATGAGAAAGCTCTTCCCT 600
Db 676 CTTTCTCTCTACAGAAATTTCAATTTACGCCCAACATCAATTCAGAAAGAGCTTTTCTCCA 735

QY 601 CAATTTGAGAGCTATCAACAGCTCTATGAAATGCTCTTCCAGGTTTGGATGGATGTT 660
Db 736 AAAATTCAGAGCAAGAAAGAACTGTATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 795

QY 661 GGTGAGCCTTAGGACTACATTTACTTGGCTTTCGCTTCTTTATCTAGACTCGATTAGGAGAAA 720
Db 796 GGCACACCAAGAGATTACATTTACTGCGCTCAGACTCTATCTGATTTCTTAAAGAAACAC 855

QY 721 TCAGTGCACCAAGTACCTTACTGACACACATGCTTCTGCGCAATCTGCTGCTGCTGCTGCTGCT 780
Db 856 TCTTACCTTAAATTTGCTTTCAGGACACACATTTGTCGGAATGTCATAGTGGATGAATCT 915

QY 781 GCCAGATTGGAGAGGTTGCTGATTTGCTGATGCTGCTGATGCTGCCATTTGACCTGGGTGTT 840
Db 916 GCCAAGATTGGAGAGGTTGTTGATGAGGACCAAGATGTTGCAATTTGCTGCTGCTGCTGCTGCT 975

QY 841 GTGAGAGGAGGCTGAGGCTTTTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
```

```
Db 976 ATTGAGTCTGGAGTTAGACTCTCCGTTGCACTGTGATGCGAGGAGTCCGATCAAGAAA 1035
QY 901 CATGCTTCCATCTCAAAACAGCATTATCGGCTGGCACTCAACTGTCTCAATGGCGACGG 960
Db 1036 CATGCATGCATCTCAGGTAGCATCATTTGGCTGGCACCTCTACTGTTGGACAATGGGCTCGT 1095
QY 961 ATAGAGAAATAGACTATCTCTGGGGGAGATGTTTCAATGTTGTGTGATGAGGTTGTACAGCAAT 1020
Db 1096 GTTGAGAACATGACATTTCTCGGGCAAGATGCTCCATCTCTTCTGATGAATTTTACAGCAAT 1155
QY 1021 GCGGCTGTTGTTCTCCACATARAAGAGATCAAGTCAAGCAATCTTGAAGCCTTGAGATCGTC 1080
Db 1156 GGAGGTGATGTTTGGCCCCACAGAGAGATCAATCCACTATATGAAACCTGAAATCGTG 1215
QY 1081 ATGTGA 1086
Db 1216 ATGTGA 1221

RESULT 12
AF022716      1478 bp      mRNA      linear      PLN 05-JAN-1999
LOCUS      Solanum tuberosum GDP-mannose pyrophosphorylase (Gmp) mRNA,
DEFINITION      complete cds.
ACCESSION      AF022716
VERSION      AF022716.1 GI:4103323
KEYWORDS
SOURCE      Solanum tuberosum.
ORGANISM      Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE      1 (bases 1 to 1478)
AUTHORS      Keller, R., Kossmann, J. and Willmitzer, L.
TITLE      Direct Submission
JOURNAL
FEATURES
source
/organism="Solanum tuberosum"
/db_xref="taxon:4113"
1..1478
/gene="Gmp"
136..1221
/gene="Gmp"
/codon_start=1
/product="GDP-mannose pyrophosphorylase"
/protein_id="AA001737.1"
/db_xref="GI:4103324"
/translation="MKALILVGGFOTRLRPLPLSVKPKPLVEFANKPMILHQIEALKAV
GTYEVLAINYOEVLNFKFEASLGKITCSQETPLGTAGPLALARDKLIDDSG
EPFVLNSDVISSEYFKEMIOFPHKSHGCEASLMVTKVDEPSKYGVVYMEESTCOVERF
VERKPLFVGNKINAGFYLLNPSVLDRIQLRPTSTIEREVFFKIAAEKKLYAMVLDPFWM
DIGOPRDYITGLRLYLDLSLAKSPHIVGNYIVDESARKIGEGCLIPDVAIG
SGCVIESGVRUSRCTVMRGVRIKKHAKICSGSIIGHWSTVGOWARVENMTILGEDVHVC
DEIYSGGVLPKHKEIKSILKPEIVM"
BASE COUNT      395 a      279 c      334 g      470 t
ORIGIN

Query Match      59.9%; Score 650.8; DB 6; Length 1478;
Best Local Similarity 75.0%; Pred. No. 1.3e-191;
Matches 814; Conservative 0; Mismatches 272; Indels 0; Gaps 0;

QY 1 ATGAGGCCCTCATCTTGTGCGGGGTTTCGGAACCGCTTCGCGCTTTGACTCTGAGC 60
Db 136 ATGAAGCCATATCCTGTTGGAGGGTTGCGGTACTGCGGTACAGGCCACACCCCTCAGT 195

QY 61 TTCCCGAAACCCCTCGTGGATTTTGCACAAAGCCCATGATTTGCACACGATCGAAGCT 120
Db 196 GTCCCAAGCCACTCGTCGAAATTTGCTAACAAACCAATGATTTTGCATCAGATTGAGGCT 255

QY 121 TTGAAAGAAAGTGGGGTCCACAGAGTGGTTTGGCTATCAACTATCCACCAAGCTGAATG 180
Db 256 CTCAGGCTGTTGGAGTAAACCAAGTTGTACTGGCTATTAACTACCACCAAGCTGAGGTGATG 315

QY 181 ATTAATTTCTTGAAGACATTTGAGGATAAGCTTTGGCATCACAAATTACATGCTCCCAAGAG 240
Db 316 CTGAAGCTTTTCAAGAAATTTGAGGAAGCCCTTGGAAATCAAGATCAAGCTGTTCTCAAGAA 375

QY 241 ACTGAGCCCTTAGGAACCGCTGCGCTCTTGTCTAGCAAGGACCAAGCTTGCAGATGGA 300
Db 376 ACTGAACCACTTGGCACTGCTGCTGCTGCTGCTTGGCTAGAGTAAGCTAGTAGATGAC 435

QY 301 TCTGCCACCAATTCCTTCTCAACAGTGATGTCATGAAGCAATACCAATTCCTGCTGAA 360
Db 436 TCTGGTGAACCATTTTCTTCTTAAAGTGATGTTATCAGTGAATATCCATTCAGGAG 495

QY 361 CTCATCAAAATTCACAAATGTCATGTTGAGGAGGCAACATGATGTCACATAGGTGGAT 420
Db 496 ATGATTCAAATTCACAAATCCCAATGCTGTTGAGGCTTCTTGTGATGTCACCAAGGTGGAT 555

QY 421 GAACCATCAAAATACGGTGTGTGGTTTATGGAGGAGGCAACTGGCAGGTGGAAAGGTTT 480
Db 556 GAGCCTTCTAAATATGCTGTTGTGTCATGGAAGCAATCCACTGCGGCAAGTAGACATTT 615

QY 481 GTTGAAGCCCAAAAATATTTCTGCGCTTAAACATCAATCTCTGCGCATTTACTTACTCAAC 540
Db 616 GTGGAGAAGCCAAAGTTATTTGTGGCAACAAGATCAATGCTGATTTTACCTGCTGAAC 675

QY 541 CCATCTGCTCTGACCGCATTTAGCTGAGGCGCAACATCAATATGAGAAAGCTCTTCCCT 600
Db 676 CTTTCTCTCTACAGAAATTTCAATTTACGCCCAACATCAATTCAGAAAGAGCTTTTCTCCA 735

QY 601 CAATTTGAGAGCTATCAACAGCTCTATGAAATGCTCTTCCAGGTTTGGATGGATGTT 660
Db 736 AAAATTCAGAGCAAGAAAGAACTGTATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 795

QY 661 GGTGAGCCTTAGGACTACATTTACTTGGCTTTCGCTTCTTTATCTAGACTCGATTAGGAGAAA 720
Db 796 GGCACACCAAGAGATTACATTTACTGCGCTCAGACTCTATCTGATTTCTTAAAGAAACAC 855

QY 721 TCAGTGCACCAAGTACCTTACTGACACACATGCTTCTGCGCAATCTGCTGCTGCTGCTGCTGCT 780
Db 856 TCTTACCTTAAATTTGCTTTCAGGACACACATTTGTCGGAATGTCATAGTGGATGAATCT 915

QY 781 GCCAGATTGGAGAGGTTGCTGATTTGCTGATGCTGCTGATGCTGCCATTTGACCTGGGTGTT 840
Db 916 GCCAAGATTGGAGAGGTTGTTGATGAGGACCAAGATGTTGCAATTTGCTGCTGCTGCTGCTGCT 975

QY 841 GTGAGAGGAGGCTGAGGCTTTTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
```

Db	256	CTCAAGGCTGTGGAGTAACCGAAGTTGTACTGGCTATTAACTACCAACCTGAGGTGATG	315
QY	181	ATTAATTTCTTGAAGGACTTTGAGGATAAGCTTGATGCATCACAAATTACATGCTCCCAAGAG	240
Db	316	CTGAACTTCTTGAAGAATTTGAGGCAAGCTTTGGAATCAAGATCACCTGTTCTCAAGAA	375
QY	241	ACTGAGCCCTTAGGAACCGCTGGCCCTCTTGCTCTAGCAAGGACAAGCTTCGGGATGA	300
Db	376	ACTGAACACCTTGGACACTGCTGGTCCCTTGCTTTGGCTAGAGATAAGCTAGATGAC	435
QY	301	TCTGCCACGACCTCTTTGTCTCAACAGTGTATGATCAATGAAGCAATACCATTTGCTGAA	360
Db	436	TCTGGTGAACCAATTTTGTCTTAAACAGTGTATGATCAAGTGAATATCCATTCAAGGAG	495
QY	361	CTCATCAAAATTTCAACAGTGTATGCTGGTGGAGCAACAATTTATGCTACTAAGTGGAT	420
Db	496	ATGATTTCAATTTCCACAATCCATCGTGTGAGGCTTCTTTGATGTTGACCAAGGTGAT	555
QY	421	GAACCATCAAAATACCGTGTGTGTTATGAGGAGGCAACTGGCAGGTTGGAAGGTTT	480
Db	556	GAGCCTTCTAATATGTTGTTGTTCTCATGGAAGAATCCACTGGGCAAGTAGAGAGATT	615
QY	481	GTGAGAGCCAAAATATTTGTGGGTAAACAGATCAATGCTGGGATTTACTTACTGAAC	540
Db	616	GTGGAGAGCCAAAGTTATTTGTTGGCAACAAGATCAATGCTGGATTTTACCTGCTGAAC	675
QY	541	CCATCTGCTCTGACCGATTGAGCTGAGGCCAACATCAATGAGAAGAGCTTTCCT	600
Db	676	CCTTCTGCTTAGACAGATTTCAATTAGGCCAACATCAATGAGAAGAGGTTTTTCCA	735
QY	601	CAAAATGCAAGTGAATCAACAGCTCTATCAATGCTGCTTCCAGGTTTTTGGATGATGT	660
Db	736	AAAATTTGAGCAGAGAAGAACTGTATCAATGCTGCTTACCTGGATTTTGGATGACATT	795
QY	661	GGTACGCTTAGGACTACATTACTGCTGCTGCTCTTATCTAGATCGATAGGAAGAA	720
Db	796	GGCCAAACCAAGAGATTACATTACTGGCCTCAGACTCTATCTGGATTCTTTAAAGAACA	855
QY	721	TCAGTGCACAGCTAGCTACTGAGCACATGTTGTTGGCAATGCTGCTGCATGAGAC	780
Db	856	TCTTCACTTAATTTGGCTTCAGGACCACATGTTGCGAAATGTCATGATGATGAATCT	915
QY	781	GCCAAATTTGAGAGAGTTGTCTGATGCTGATGCTGCATGCGCACTTGCACTGGGTGT	840
Db	916	GCCAAATTTGAGAGAGTTGTTGATGATGACAGATGTTGCAATTTGGTCTGTTGTG	975
QY	841	GTGAGGAGCGGTGAGGCTTTCCGCTGCACTGCTGCTGCGCGGCTGCGTATCAAGAAG	900
Db	976	ATTGAGTCTGGAGTTAGACTCTCCGTTGCACTGTGATGCGAGGAGTCCGCATCAAGAA	1035
QY	901	CATGCTGCATCTCAACAGCATTTATCGCTGGCACTCAACTGTTGGTCAATGGGCACG	960
Db	1036	CATGCTGCATCTCAGGTAGCATCATTTGGTGGCACTCTACTGTTGGCAATGGGCTCGT	1095
QY	961	ATAGAGAAATGACTATCTCTGGGGAGGATGTTCAATGCTGCTGATGAGGTGACAGCAAT	1020
Db	1096	GTTGAGACATGACCATTTCTCGGGAAGATGTCATGTTGTGATGAATTTACAGCAAT	1155
QY	1021	GGCGGTGTTGTTCTCCCAATAAAGAGATCAAGTCAAGCATTTCTGAAGCCTGAGATCGPC	1080
Db	1156	GGAGGTGATGTTTGGCCCAAGAGATCAAAATCCAGTATATTGAAACCTGAAATCGTG	1215
QY	1081	ATGTGA 1086	
Db	1216	ATGTGA 1221	
RESULT 13			
AB066279			
LOCUS	AB066279	1471 bp	mRNA linear
DEFINITION	Nicotiana tabacum GMPase mRNA for GDP-D-mannose pyrophosphorylase, complete cds.		
ACCESSION	AB066279		

VERSION	AB066279.1	GI:14971012	
KEYWORDS	Nicotiana tabacum cell_line:BY-2 cDNA to mRNA.		
SOURCE	Nicotiana tabacum		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; eusterids I; Solanales; Solanaceae; Nicotiana.		
REFERENCE	1 Tabata,K., Takaoka,T., Tateishi,Y. and Esaka,M. Gene expression of ascorbic acid-biosynthetic and -oxidizing enzymes and ascorbic acid content in tobacco		
AUTHORS	Unpublished		
TITLE	2 (bases 1 to 1471)		
JOURNAL	Esaka,M., Takaoka,T., Tateishi,Y. and Tabata,K.		
REFERENCE	Submitted (18-JUL-2001) Muneharu Esaka, Hiroshima University,		
AUTHORS	Faculty of Applied Biological Science, Kagamiyama,		
TITLE	Higashi-Hiroshima, Hiroshima 739-8528, Japan		
JOURNAL	(E-mail:mesaka@hiroshima-u.ac.jp, Tel:81-824-24-7927, Fax:81-824-24-7927)		
FEATURES	Location/Qualifiers		
Source	1..1471		
	/organism="Nicotiana tabacum"		
	/db_xref="taxon:4097"		
	/cell_line="BY-2"		
	/note="cultured"		
gene	1..1471		
	/gene="GMPase"		
CDS	220..1305		
	/gene="GMPase"		
	/EC_number="2.7.7.22"		
	/function="ascorbic acid biosynthetic enzyme/carbohydrate biosynthetic enzyme"		
	/codon_start=1		
	/product="GDP-D-mannose pyrophosphorylase"		
	/protein_id="BAB62108.1"		
	/db_xref="GI:14971013"		
	/translation="MKALILVGGFGRLRLTLSPKPLVPEFANKPMILHOIEALKAV GTYEVVLAIFYOPEEMLNELKFEANLGIKTCSETEPLGTAGPLALARDKLIDDSG RFFVFLNSDVISSEYPKEMIAFHKSHGGASLMTVDPSKYGVVYMERSTQGVRF VBKPLFVGNKINAGLYLLNPSVLDRIQRPTSEIEKVEFPKIAAEKFLYAMVLPFGWM DTGQPRDYTPGLRLYLDLSKNSPKLASGSHVGNVIDESAKIEGCLIGDPDAIG SGCVIESGVRLSRCTVMGVRIKKHACISSIIGWHSTVGQWARVENMTILGSDVHC DELYSNGGVVLPHKEIKSLKPEIVM"		
BASE COUNT	407 a	268 c	350 g 446 t
ORIGIN			
	Query Match 58.7%; Score 638; DB 8; Length 1471;		
	Best Local Similarity 74.2%; Pred. No. 1.3e-187;		
	Matches 806; Conservative 0; Mismatches 280; Indels 0; Gaps 0;		
QY	1	ATGAAGCCCTCATCTTGTTCGGGGTTCGGAACCGCCTTCGGCCCTTTGACTCTGAGC	60
Db	220	ATGAAGCGGTTATTCTTGTGGAGGGTTCGGTACTCGTTGAGGCCATTGACGCTCAGT	279
QY	61	TTCCCGAAACCCCTCGTGGATTTGCAACAAGCCCATGATCTGCACAGATCGAAGCT	120
Db	280	GTCCCAAGGCATAGTTGAATTTGCCAACAACCCCATGATTTGCATCAGATTGAGGCT	339
QY	121	TTGAAAGAAGTTGGGTTCACAGAGTGGTTTGGCTATCAACTATCCCGCAGAGGTAATG	180
Db	340	CTCAAGCAGTCGAGTAAGTGAAGTGTACTGGCTATTATTAACCAACCTGAGGAGATG	399
QY	181	ATTAATTTCTTGAAGGACTTTGAGGATAAGCTTGGCATCAAAATTAACATGCTCCCAAGAG	240
Db	400	CTGAACCTTCTTGAAGAATTTGAGGCAAACTCTGGAATCAAGATCACCTGTTCTCAAGAA	459
QY	241	ACTGAGCCCTTAGAACCCTGGCCCTCTTCTCTACACAGGACAGCTTGGCGGATGGA	300
Db	460	ACTGAACCACTTGGCACTGCAAGGTCCCTTGTGCTTAGAGATAAGCTGATTGATGAC	519
QY	301	TCCTGGCCAGCCATCTCTTGTGCTTCAACAGTGATGCTCATAGCGGAATACCCATTTGCTGAA	360





DB	32691	TAGGTACACAGATCAAGCGTGGGATTTATCTTCTGAACCCATCTGTTCTTGATAAGATTG	32750
QY	563	AGCTAGGCGCAAGATCAATAGAAAGAGAGTGCTTGGCTCAAAATGGAGCTGATCAACAGC	622
DB	32751	ACGTAACAGCGGACTTCAATCGMAAAAGACAGACTTTCCTAAAGATTGCAGCAGCCCAAGGC	32810
QY	623	TCTATGCAATGGTCTTCCAGGTTTGGATGATGATGCTGGTACGCCVAGGCGCATACATTA	682
DB	32811	TCTATGCTATGTTGCTTACCAGGGTTTGGATGGACATTTGGGCAACCCCGTGACTACATAA	32870
QY	683	CTGCGCTTGGCTTATCTATGACTCGATAGGAAGAAATCAGCTGCCACCTAGCTACTG	742
DB	32871	CGGTTTGAAGACTCTACTTAGACTCCCTTAGGAAGAAATCTCTGCCAAATTAAACAGT	32930
QY	743	GAGCAGATGTTGTGCAATGTGCTGGTGCATGAGAGCGCAAGATTGGAGAGGTTGTC	802
DB	32931	GGCCACACATAGTTGGAAATGTTCTGTTTCACGAACCCGCTACAATTGGGAGGATGTT	32990
QY	803	TGATTGGTCTGATGTCGCATTTGACCTGGGTGTTGTTGGAGGAGCGGCTGAGGCTTT	862
DB	32991	TGATTGGACAGACAGCTTGCCATTTGTCAGGCTGCTGATTTGATGAGGAGTCACTCT	33050
QY	863	CCGCTGCACTGTCATCGCGGCGTGGTATCAAGAGCATGCTTCATCTCAACAGCA	922
DB	33051	CCGATGACAGGTCATCGCTGGAGTCCGATCAAGAGCATGCTGTATCTCAGCAGTA	33110
QY	923	TATCGCTGGCACTCAACTGTTGGTCAATGGCAGCGATAGAGATATGACTATCTCTG	982
DB	33111	TCATCGCTGGCACTCAACGGTGGTCAATGGCCAGGATCGACACATGACCATCTCTG	33170
QY	983	GGGAGATGTTGATGTTGTGATGAGGTGTACAGCAATGGCGGTGTTGTTCTCCACATA	1042
DB	33171	GTGAGGATGTTGATGAGGATGAGATCTATAGCAATGGAGGATGTTTTCACACACA	33230
QY	1043	ACAGATCAACATCAACCATTCGAACCCGACATGCTCATCTCA	1086
DB	33231	AGGATCAAAATCAACATCTTTGAAGCCAGAGATAGTGATGA	33274
RESULT	15		
ATT22E16			
LOCUS	ATT22E16	103240 bp	DNA linear PLN 23-FEB-2000
DEFINITION	Arabidopsis thaliana DNA chromosome 3, BAC clone T22E16.		
ACCESSION	AL132975		
VERSION	AL132975.1	GI:6434228	
KEYWORDS	Arabidopsis thaliana.		
SOURCE	Arabidopsis thaliana		
ORGANISM	Arabidopsis thaliana		
REFERENCE			
AUTHORS	Benes,V., Wurmbach,E., Drzonek,H., Ansorge,W., Mewes,H.W., Lemcke,K., Mayer,K.F.X., Quetier,F. and Salanoubat,M.		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 103240)		
AUTHORS	EU Arabidopsis sequencing,project.		
TITLE	Direct submission		
JOURNAL	Submitted (23-FEB-2000) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: lemcke@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de Project Coordinator: Marcel Salanoubat and Francis Quetier, Grouperment d'Interet Public, Centre National de Sequencage - GENOSCOPE, 2 rue Gaston Cromieux, BP191, 91006 Evry Cedex, France;		
COMMENT	http://www.genoscope.cns.fr Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/tha/.		
FEATURES	Location/Qualifiers		
source	1..103240		
	/organism="Arabidopsis thaliana"		
	/variety="Columbia"		
	/db_xref="taxon:3702"		

misc_feature	1..8852	/chromosome="3"	
gene	992..3164	/note="Overlap to BAC T26112 (EMBL:AL132954)"	
		/gene="T22E16.10"	
CDS	101n(992..1181,2134..3164)		
		/gene="T22E16.10"	
		/note="similarity to predicted proteins, Arabidopsis thaliana"	
		/codon_start=1	
		/product="putative protein"	
		/protein_id="CAB75893.1"	
		/db_xref="GI:7076778"	
		/translation="MGPIKTIKKKRAEKVDRNVLLAATAAATASASAAAALNNDDDDSSSSOSLMDWDFSRRIYGGSDPKTFESVFKISRKTFDYICSLVADTFKAPANESDSNGNLSLNDRAVALRIGSGESLSVIGETFGMNOSTVSOITWRVSEMERAIHLSWPKLDEIKSFEKISGLPNCAGADITHIYNLPAVSPKRWLDGKFNSTLQAVYDPMRLDVIAGWPGSUNDVYLNKSGFKLVKRGKLNKGLKPLSEKREYIIVGDSGFPDLLPWLTPYCKPTSLPOTEFNKRHSATKAQMAISKLKDRWRIINGVMMPDRNLRPRIIFVCCLLHNTIIDMEDQTLDDQLSQQHDHMYRQSRKLADEASSVLRDELSDLQCGKNSSA"	
exon	992..1181	/gene="T22E16.10"	
		/number=1	
intron	1182..2133	/gene="T22E16.10"	
		/number=1	
exon	2134..3164	/gene="T22E16.10"	
		/number=2	
gene	3660..5330	/gene="T22E16.20"	
CDS	complement(join(3660..3728,3821..4575,5038..5077,5262..5330))		
		/note="strong similarity to synaptic glycoprotein SC2 spliced variant, Homo sapiens, EMBL:AF038958"	
		/codon_start=1	
		/product="glycoprotein-like"	
		/protein_id="CAB75894.1"	
		/db_xref="GI:7076779"	
		/translation="MKVTVVSRSGRVILKAPLDLPDSATVADLQEAPEHRAKRPSPSRQRUTLPVTGSKDKPVLNLSKSKSEYCDGNNSLTIVFKDLGAQSYRTLFFFEYILGPLLIYFVYFYFYKFGYGEDCVIHPVQTYAMY YC FHYFKRIETFFVHRFSHATS PIGNVFNCAVYVSGFAYIAYVNHPLYTPVSDLOMKRIGFGGLVCCOVANFYCHILK NLRDPSAGGYQLPRGFLENITVCANTTEIYQWLGNIATQITAGVYFLAVALIINT NWALGKHSRLRKLEIDGDKPKYPRRWVILPPEL"	
exon	complement(3660..3728)		
		/number=1	
intron	complement(3729..3820)		
		/number=1	
exon	complement(3821..4575)		
		/number=2	
intron	complement(4576..5037)		
		/number=2	
exon	complement(5038..5077)		
		/number=3	
intron	complement(5078..5261)		
		/number=3	
exon	complement(5262..5330)		
		/number=4	
gene	10334..11350	/gene="T22E16.30"	
		/gene="T22E16.30"	
CDS	join(10334..11042,11136..11350)		
		/note="EMBL:AF155818 most likely is not full length, therefore N-terminal stretch"	
		/codon_start=1	
		/product="zinc finger protein OBP3"	
		/protein_id="CAB75895.1"	
		/db_xref="GI:7076780"	
		/translation="MOOGNHOLECVTTTONPNNYIROLSPPTSOVAGSSOARVNSM VFRATARKVPLPEALNCPDSCNTKFCVFNYSLTOPHFKTCRRVTRGSGSLN VPVGGFRNKRKSKSKSTVYVYSTONTTSTSSLSRPSYSNFSKFSYQGIPEFNSN	







GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2002, 02:04:11 ; Search time 255 Seconds  
(without alignments)  
9590.863 Million cell updates/sec

Title: US-09-374-967-1  
Perfect score: 1086  
Sequence: 1 atgaagccctcattctgt.....agcctgagatgctcatgtga 1086

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002.\*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*
- 24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1010	93.0	1481	21 AAC44043	Zea mays DNA fragm
2	690.8	63.6	1109	21 AA298316	A. thaliana gene i
3	690.8	63.6	1474	21 AAC48703	Arabidopsis thalia
4	690.8	63.6	1480	21 AAC38731	Arabidopsis thalia
5	690.8	59.9	1478	20 AAX77618	Potato GDP-mannose
6	443	40.8	556	21 AAC39741	Zea mays DNA fragm
7	417.2	38.4	1358	23 ABL20809	Drosophila melanog
8	417.2	38.4	1476	23 ABL02105	Drosophila melanog
9	399.6	36.8	1512	21 AAC75966	Human ORFX ORF1521

10	399.6	36.8	1559	21	AA294202	Human transferase
11	358	33.0	1092	22	ABA01227	NDP-hexose pyropho
12	357.4	32.9	3220	22	AAH18598	Human cDNA sequenc
13	322	29.7	990	22	AA41146	cDNA encoding nove
14	300.4	27.7	1351	21	AAF13173	Aspergillus oryzae
15	262.6	24.2	458	21	AAC43567	Zea mays DNA fragm
16	259.4	23.9	487	21	AAC41437	Zea mays DNA fragm
17	221	20.3	467	21	AAC41446	Zea mays DNA fragm
18	215.6	19.9	4313	23	ABL02104	Drosophila melanog
19	215.6	19.9	4334	23	ABL20808	Drosophila melanog
20	187	17.2	212	24	ABL72790	Corn tassal-derive
21	168.8	15.5	802	21	AAO2132	Human colon cancer
22	162.6	15.0	625	21	AAF07767	Fusarium venenatum
23	141.2	13.0	243	24	ABN19689	Human ORFX polynuc
24	124	11.4	3377	23	ABL02088	Drosophila melanog
25	119.8	11.0	300	21	AAA00584	Human colon cancer
26	107	9.9	670	22	AA41147	cDNA encoding nove
27	92.4	8.5	514	22	AAH07288	Human cDNA clone (
28	85.2	7.8	4297	22	AA541953	Genomic sequence #
29	83.2	7.7	700	21	AA51834	Arabidopsis thalia
30	80.4	7.4	388	22	AAF6902	Novel human polynu
31	74.2	6.8	787	22	AAH03616	Human cDNA clone (
32	74.2	6.8	1520	22	AAH13791	Human cDNA sequenc
33	74.2	6.8	1535	20	AA87904	Human cell cycle r
34	74.2	6.8	1550	22	AA541153	cDNA encoding nove
35	74.2	6.8	1822	22	AAH15813	Human cDNA sequenc
36	73.2	6.7	787	24	AAD23700	Soybean Bng12D ma
37	63	5.8	1248	21	AAC45822	Arabidopsis thalia
38	61.4	5.7	1804	21	AAC37770	Arabidopsis thalia
39	58.4	5.4	7902	19	AAV52245	Streptococcus pneu
40	57.6	5.3	34980	22	AAH41224	Pyrococcus abyssi
41	57.6	5.3	34980	22	AAH41225	Pyrococcus abyssi
42	56	5.2	1275	21	AAC42394	Arabidopsis thalia
43	55.2	5.1	282	24	ABL74497	Corn tassal-derive
44	55.2	5.1	690	19	AAV33009	Nucleotide sequenc
45	53.2	4.9	509	22	AAI85712	Human polynucleoti

ALIGNMENTS

RESULT 1  
AAC44043  
ID AAC44043 standard; DNA; 1481 BP.  
XX AAC44043;  
AC AAC44043;  
XX  
XX 18-OCT-2000 (first entry)  
DT  
DE Zea mays DNA fragment SEQ ID NO: 41409.  
XX  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic;  
KW pathway; promoter; termination sequence; corn; ss.  
XX  
OS Zea mays subsp. mays.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139730.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140699.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0141854.  
PR 02-JUL-1999; 99US-0142035.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 18-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 28-JUL-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 08-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149930.  
PR 23-AUG-1999; 99US-0149902.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 28-SEP-1999; 99US-0156596.  
PR 29-SEP-1999; 99US-0157117.  
PR 04-OCT-1999; 99US-0157753.  
PR 05-OCT-1999; 99US-0157953.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158569.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.

[illegible]



PR	28-MAY-1999;	99US-0136782.	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	99US-0137222.	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	99US-0137528.	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	99US-0137502.	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	99US-0137724.	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	99US-0138094.	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	99US-0138540.	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	99US-0138847.	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	99US-0139119.	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	99US-0139452.	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	99US-0139453.	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	99US-0139492.	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	99US-0139454.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139462.	PR	25-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139750.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151065.
PR	21-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	22-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151080.
PR	23-JUN-1999;	99US-0139899.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140698.	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0142154.	PR	20-SEP-1999;	99US-0154779.
PR	02-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155139.
PR	06-JUL-1999;	99US-0142390.	PR	23-SEP-1999;	99US-0155486.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-0155659.
PR	09-JUL-1999;	99US-0142920.	PR	28-SEP-1999;	99US-0156458.
PR	12-JUL-1999;	99US-0142977.	PR	29-SEP-1999;	99US-0156596.
PR	13-JUL-1999;	99US-0143542.	PR	04-OCT-1999;	99US-0157117.
PR	14-JUL-1999;	99US-0143624.	PR	05-OCT-1999;	99US-0157753.
PR	15-JUL-1999;	99US-0144005.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144085.	PR	07-OCT-1999;	99US-0158029.
PR	16-JUL-1999;	99US-0144086.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144325.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144325.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144332.	PR	14-OCT-1999;	99US-0159330.
PR	19-JUL-1999;	99US-0144333.	PR	14-OCT-1999;	99US-0159330.
PR	19-JUL-1999;	99US-0144333.	PR	14-OCT-1999;	99US-0159329.
PR	19-JUL-1999;	99US-0144334.	PR	18-OCT-1999;	99US-0159584.
PR	20-JUL-1999;	99US-0144335.	PR	21-OCT-1999;	99US-0160741.
PR	20-JUL-1999;	99US-0144352.	PR	21-OCT-1999;	99US-0160767.
PR	20-JUL-1999;	99US-0144632.	PR	21-OCT-1999;	99US-0160768.
PR	20-JUL-1999;	99US-0144884.	PR	21-OCT-1999;	99US-0160770.
PR	21-JUL-1999;	99US-0144884.	PR	21-OCT-1999;	99US-0160814.
PR	21-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160815.
PR	21-JUL-1999;	99US-0145086.	PR	22-OCT-1999;	99US-0160980.
PR	21-JUL-1999;	99US-0145088.	PR	22-OCT-1999;	99US-0160981.
PR	22-JUL-1999;	99US-0145085.	PR	22-OCT-1999;	99US-0160981.
PR	22-JUL-1999;	99US-0145087.	PR	22-OCT-1999;	99US-0160989.
PR	22-JUL-1999;	99US-0145087.	PR	22-OCT-1999;	99US-0160989.
PR	22-JUL-1999;	99US-0145192.	PR	25-OCT-1999;	99US-0161404.
PR	23-JUL-1999;	99US-0145145.	PR	25-OCT-1999;	99US-0161405.
PR	23-JUL-1999;	99US-0145218.	PR	25-OCT-1999;	99US-0161406.
PR	23-JUL-1999;	99US-0145224.	PR	26-OCT-1999;	99US-0161359.
PR	26-JUL-1999;	99US-0145276.	PR	26-OCT-1999;	99US-0161360.
PR	27-JUL-1999;	99US-0145913.	PR	26-OCT-1999;	99US-0161361.
PR	27-JUL-1999;	99US-0145918.	PR	28-OCT-1999;	99US-0161920.
PR	27-JUL-1999;	99US-0145919.	PR	28-OCT-1999;	99US-0161920.
PR	28-JUL-1999;	99US-0145919.	PR	28-OCT-1999;	99US-0161993.
PR	28-JUL-1999;	99US-0145951.	PR	29-OCT-1999;	99US-0162142.
PR	02-AUG-1999;	99US-0146380.			
PR	02-AUG-1999;	99US-0146388.			
PR	02-AUG-1999;	99US-0146389.			
PR	03-AUG-1999;	99US-0147038.			
PR	04-AUG-1999;	99US-0147204.			



Query Match	63.6%	Score	690.8	Db	21	Length	1474
Best Local Similarity	77.3%	Pred.	No. 2.7e-210				
Matches	839	Conservative	0	Mismatches	247	Indels	0
QY	1	ATGAGGCGCCATCTTCTTGTCGGGGGTTTCGGAAACCGCCTTCGGCCCTTGACTCTGAGC	60				
Db	155	ATGAAGGCACTCATCTTGTTGTGAGGCTTCGGCACTCGCTTGAGACCAATGACTCTCAAT	214				
QY	61	TTCCGGAACCCCTCGCTGCAATTTTTCGAAACACCCCAATTCGACACAGATCCAAAGCT	120				
Db	215	TTCCCAAGCCCCCTTGTTGATTTTGCTAATAAACCCATGATCCTTCATCAGATAGGCGT	274				
QY	121	TTGAAGAAGGTTGGGTCACAGAGGTGGTTTTGGCTATCAACTATCGCCACAGAGTAATG	180				
Db	275	CTTAAAGCACTTGGAGCTTGATGAAGCTGCTTTGGCCATCAATATACGCCAGAGGTGATG	334				
QY	181	ATTAATTTCTTGAGGACCTTGAGGATAGCTTGGCATCACAATTTACATGCTCCCAAGAG	240				
Db	335	CTGAACCTCTTGAAGGACTTTGAGACCAAGCTGGAAATCAAAATCACTTGCTGCACAAAG	394				
QY	241	ACTGAGGCGTTAGGAACCGGTCGCCCTTCGTCGTAAGCAAGCAAGCGTTGCCGATCGCA	300				
Db	395	ACGAGCCACTAGGTACCGCTGGTCTCTGGCTCTAGCGAGGAGCAAAATGCTTGATGGA	454				
QY	301	TCTGGCCAGCACTCTTGTCTCCACACAGCTGATGTCATAAGCGAAATACCAATTTGCTCAA	360				
Db	455	TCTGGAGAGCCCTTCCTTCTTAACAGTGATGTCATTTAGTACCTTCTTAAGAA	514				
QY	361	CTCATCAAAATTTACAAGTGTCAIGTGGTGAGGCAACAATTTATGGTCACTAAGTGGAAT	420				
Db	515	ATGCTTGAGTTTCACAAATCTCACGGTGGGGAAGCGCTCCATTAATGGTAAACAAGCTGAT	574				
QY	421	GAACCATCAAAATACGCTGTCGCTTATGAGGAGGCAACTCGACGGTGGCAAGGTTT	480				
Db	575	GRACCGTCGAATATGGAGTGGTTGTTATGGAAGAAAGCACCTGGGAAGAGTGGGAAGTTT	634				
QY	481	GTTGAGAAGCCAAAATATTTGTGGGTAAACAAGATCAATGCTGGGATTTACTATCTCAAC	540				
Db	635	GTGGAAGGCCAAAACCTGATGTAGGTAACAAGATCAAGCTGGGATTTATCTTCTGAC	694				
QY	541	CCATCTGCTCTTGACCGCATTTAGCTTGAGGCCCAACATCAATTTGAGAAAGAGCTTCCCT	600				
Db	695	CCATCTGCTCTTGATTAAGATTTGAGCTTAACCGCATTTCAATTCGAAAAACAGACTTCCCT	754				
QY	601	CAAAATTCAGCTCATCAACAGCTCATGCAATGGTCTCTCCAGGTTTTTGGATGATGTT	660				
Db	755	AGATTTGACAGCGCAAGGCTCTATGCTATGGTGTACCAAGGTTTGGATGGACATTT	814				
QY	661	GCATCAGCCTAGGACATACATTTACTGGCTTTCCTTTCTTTATCTAGACTCGATTAGGAAGAA	720				
Db	815	GGCAACCCCGTGACTTACATAACGGGTTTGGAGCTCTACTTAGACTCCCTTAGGAAGAA	874				
QY	721	TCAGCTGCCAAGCTAGCTACTTGAGGACATGTTGTGGCAATGTCTGGTCATCGAGAGC	780				
Db	875	TCCTCTGCCAAAATAACCAAGTGGGCAACACATAGTTGGCAATGTTCTTTGTCACCAAAACC	934				
QY	781	GCCAAAGATTTGGAGAGGTTGTCTGATTTGGTCTCTGATGTGCCATTTGGACCTTGGGTGTT	840				
Db	935	GCTACAAATTTGGGAAGGATGTTTGATTTGACAGACAGTGTGCCATTTGGTCCGAGGCTGCAT	994				
QY	841	GTGGAGCAGCGCTGAGGCTTTCCGCTGCATCTGTATGCGCGGCTGCGCTATCAAGAG	900				
Db	995	GTTTGAATCAGGATCAGATCTCCCGATGCAGGGTCATGCGGTGGATCTCCGATCAAGAG	1054				
QY	901	CATGCTTGCACTCAAAACAGCAATTAACGGCTGGCACTCAACTTTGGTCAATGGGCAAGG	960				
Db	1055	CATGCGTGATATCTCGAGCAGTATCATCGGGTGGCACTCAACGGTTTGGTCAATGGGCCAGG	1114				
QY	961	ATAGAGAAATGACTATCTCGGGGAGGATGTTTCATGTGTGTATGAGGTTACAGCAAT	1020				
Db	1115	ATCGAAGCAATCAGATCTCGGTGAGGATGTTTCATGTGCGCATGAGATCATATGCAAT	1174				



QY 61 TTCCCGAAGCCCTCTGCGATTTTGCACAAAGCCCAAGATTTCTGCAACAGATPCGAAGCT 120  
DB 215 TTCCCAAGCCCTTGTGATTTTGTCTAATAAACCCATGATCTTCTCATCAGATAGAGCT 274  
QY 121 TTCAAGAACTTTCGGCTCACAGAGTGGTTTGGCTATCAACTATCGCCAGAGCTAATG 180  
DB 275 CTTAGGCAGTTGGAGTTGATGAAGTGGTTTGGCCATCAATTTATCAGCCAGAGTGCAT 334  
QY 181 ATTAATTTCTTGAAGACTTTGAGGATAAGCTTGGCATCACAAATACATGCTCCCAAGAG 240  
DB 335 CTCGAATCTTGNAGACATTTGAGACCAAGCTGGAATCAAAATCACTGCTCACAAGAG 394  
QY 241 ACTGAGCCCTTAGGAACCCCTGCCCCCTTGTCTTACCAAGGACAAAGCTTGGCGATGGA 300  
DB 395 ACCGAGCCACTAGTACCCCTGCTCTGGCTCTAGCGAGAGACAAATTTGCTTGATGA 454  
QY 301 TCTGCCCCACCAATCTTCTCTCTCAACAGTATGTCATGAAGCGAATACCGATTTGCTGAA 360  
DB 455 TCTGGAGACCCCTTCTTCTTAAACAGTATGTCATTTAGTACCTACCTCTTAAAGAA 514  
QY 361 CTCATCAAAATTCACAAGTGTATGTTGATGAGGCAACAATATGTCATCAAGSTGGAT 420  
DB 515 ATGCTTGACCTTTCACAAATTCACGCTGGGAAAGCTCCATTAATGTTAAACAAAGSTGGAT 574  
QY 421 GAACCATCAAAATACGGTGTGTTGTTATGGAGGCGCAACTGGCAGGCTGCAAAAGCTTT 480  
DB 575 GAACGTCGAATATGGAGTGTGTTATGGAAGAACACCTGGGAAGAGTGGGAAGTTT 634  
QY 481 GTTGAGAGCCAAATAATTTGTGGGTAAACAAGTCAATGCTGGGATTTACTTACTGAAC 540  
DB 635 GTGGAAGGCCAAACTGTATGAGTAAACAAGATCAACGCTGGGATTTTATCTTCTGAAC 694  
QY 541 CCATGCTGCTTGACGCAATTCAGCTGAGCGCAACATCAATTGAGAAAGAGTCTTCCCT 600  
DB 695 CCATCTGTTCTTGAATGAAGATTGAGCTAAGCCGATTCATCGAAAGAGAGACTTTCCCT 754  
QY 601 CAAATTTGACGTGATCAACAGCTCTATGCAATGGTCTTCCAGTTTGTGGATGGATGT 660  
DB 755 AAGATTGCAAGCAGCAAGGGCTCTATGCTATGGCTACAGGGTTTGGATGGACAT 814  
QY 661 GGTGAGCTTAGGAGTACATTTACTGCTTTCCTGCTTTATCTAGATPCGATTTAGGAAGAA 720  
DB 815 GGGCAACCCCGTGTACTACATAACGGTTTGTAGACTCTACTTAGACTCCCTTAGGAAGAA 874  
QY 721 TCAGCTGCAACCTAGCTACTGACACATGTTGTGGCAATGCTGCTGTCATGAGAC 780  
DB 875 TCTCTGCAAAATTAACCAAGTGGGCCACACATAGTTGGGAATGTTCTTGTGACGAAAC 934  
QY 781 GCGAAGATTGGAGAAGTTGTTGATGTTGCTGATGTCGCAATTTGACCTTGGGTTGTT 840  
DB 935 GCTACAATTTGGGAAGGATGTTGATTTGATGGACACAGACCTTGGCCATGGTCCAGGCTCAT 994  
QY 841 GTGAGGAGCGGCTGAGGCTTTCCCGCTGCACTGTGATGCGCGGCTGCTATCAAGAAG 900  
DB 995 GTTGAATCAGGAGTCAAGACTCTCCCGATGACGGTTCATGCTGGAGTCCGCAATCAAGAG 1054  
QY 901 CATGCTTGCATGCTAATAAAGCAATTTGCGCTGCGCACTCAACTGTTGCTCAATGGCACGG 960  
DB 1035 CATGCTGTATCTCAGCAGTATCATCGGGTGGCACTCAACGGTTTGGTCAATGGGCCAGG 1114  
QY 961 ATAGAAATATGACTATCTGGGGAGAGATGTTCTATGTTGTGATGAGTGTAGAGCAAT 1020  
DB 1115 ATCAGACATACAGCATCTCTGCTGAGGATGTTCTATGTCAGCGCATGAGATCATGCAAT 1174  
QY 1021 GCGGTTGTTTCTCCACATAAAGAGATCAAGTCAAGCATTTCTGAAGCCTTGAGATCGTC 1080  
DB 1175 GAGAGCTTGTGTTTCCACACAGAGAGATCAAAATCAACATCTTGAAGCCAGAGATG 1234  
QY 1081 ATGTGA 1086  
DB 1235 ATGTGA 1240

RESULT 5  
AAAX77618  
ID AAX77618 standard; cDNA to mRNA; 1478 BP.  
XX  
AC AAX77618;  
XX  
DT 13-AUG-1999 (first entry)  
XX  
DE Potato GDP-mannose pyrophosphorylase cDNA.  
XX  
KW GDP-mannose pyrophosphorylase; transgenic plant; potato; plant cell;  
XX  
OS Solanum tuberosum.  
XX  
PH Key Location/Qualifiers  
FT CDS 136..1221  
FT /\*tag= a  
FT /product= "GDP-mannose pyrophosphorylase"  
XX  
PN DE19741375-A1.  
XX  
PD 01-APR-1999.  
XX  
PF 19-SEP-1997; 97DE-1041375.  
XX  
PR 19-SEP-1997; 97DE-1041375.  
XX  
PA (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.  
XX  
PI Keller R. Kossmann J;  
XX  
XX WPI: 1999-Z15846/19.  
DR P-PSDB; AAY08879.  
XX  
PT DNA encoding potato GDP-mannose pyrophosphorylase protein - and  
PT corresponding protein, antibody, transgenic plants with decreased or  
PT increased GDP-mannose pyrophosphorylase activity, etc.  
XX  
PS Claim 2b: Page 11-13: 22pp; German.  
XX  
CC This invention describes a method for the production of a novel  
CC transgenic plant containing plant cells that have been genetically  
CC modified to decrease their endogenous GDP-mannose pyrophosphorylase  
CC (GMPase) activity. The invention also describes a method for producing  
CC transgenic plants whose aerial parts mature, ripen and die earlier and  
CC more rapidly than those of corresponding wild-type plants, comprising  
CC genetically modifying a plant cell by introducing a foreign DNA molecule  
CC to decrease the cell's endogenous GMPase activity, regenerating a plant  
CC from the cell, and optionally producing further plants from this plant.  
XX  
SQ Sequence 1478 BP; 395 A; 279 C; 334 G; 470 T; 0 other;

Query Match 59.9%; Score 650.0; DB 20; Length 1476;  
Best Local Similarity 75.0%; Pred. NO. 1.7e-197;  
Matches 814; Conservative 0; Mismatches 272; Indels 0; Gaps 0:  
QY 1 ATGAAGCCCTCATTTCTTGTGGGGTTTCGGAACCCGCTTCGGCTTTTGACTCTGAGC 60  
DB 136 ATGAAGCCATTTATCTTGTGGAGGTTTCGGTACTCGGCTCAGGCACCTCAGT 195  
QY 61 TTCCCCAAACCCCTCTGGGATTTTGCACAAAGCCCATGATTTGCAACCATGCAAGCT 120  
DB 196 GTCCCAAGCCAGCTCTCGAAATTTGCTAACAACCAATGATTTTGCATCAGATTGAGCT 255  
QY 121 TTCAAAAGAACTTGGCTGCACAGAGTGGTTTGGCTATCAACTATGCCAGAGCTAATG 180  
DB 256 CTCAGGCTGTGGAGTACCGAAGTTGTACTGGCTATTAACTACCAACCTGAGGTGATG 315  
QY 181 ATTAATTTCTTGAAGGACTTTGAGGATAGGCTTGGCATCAAAATACATGCTCCCAAGAG 240  
DB 316 CTCGAATCTTGTGAAGAAATTTGAGGCAAGCTTTGGAATCAAGATCAGCTCTTCTCAAG 375

Qy	241	ACTGAGCCCTTAGGAACCGCTGGCCCTCTTGCTGTAGCAAGGACAAGCTTGGGGATGGA	300
Db	376	ACTGAACCACTTGGCACTGCTGGTCCCTTGCTTGGCTAGAGATAAGCTGATAGATGAC	435
Qy	301	TCTGCCACGCCATCTTTGTCTCAACAGTGTATGTCATAAGCGAATACCACTTTCGTGA	360
Db	436	TCTGTGAACCATTTTTTGTCTTAAACAGTGTATGTCAGTGAATATCCATTCGAAGAG	495
Qy	361	CTCATCAAAATTTACAAGTGCATGGTGTGAGGCAACAATATATGCTACTAAGSTGGAT	420
Db	496	ATGATTCATTCACAAATCCCATGGTGTGAGGCTTCTTGTATGGTGACCAAGSTGGAT	555
Qy	421	GAACCATCAAAATACGGTGTGTGTATGAGGAGGCAACTGGCAGGCTGGAAAGGTTT	480
Db	556	GAGCCTTCTAAATATGGTGTGTGTTCATGGAAGAATCCACTGGGCAAGTAGAGATTT	615
Qy	481	GTTGAGAGCCAAAATATTTGTGGGTACACAGATCAATGCTGGGATTTACTTTACTGAAC	540
Db	616	GTGGAGAGCCAAAGTTATTTGTGGCAACAAGATCAATGCTGGATTTTACCTGCTGAAC	675
Qy	541	CCATCTGTCTTGACCGCATTCAGCTGAGGCCAACATCAATTGAGAAAGAGTCTTCCCT	600
Db	676	CCTTCTGTCTAGACAGATTCATTAATACGCCCAACATCAATTGAGAAAGAGTTTTCOA	735
Qy	601	CAAAATGCACTGATCAACAGCTCTATGCAATGGTCCCTTCCAGGTTTTTGGATGGATGT	660
Db	736	AAAATGCAAGAGAAAGAACTGATGCAATGGTGTCTACCTGGATTTTGGATGGACATT	795
Qy	661	GGTCAGCCTAGGACTACATACTGCTGGTGTGCTCTTTATCTAGACTCGATTAGGAAGAA	720
Db	796	GGCCAACCAAGAGATACATTAATGCTGCTTCAGCCTTCAGACTCTATCTGGATTCCTTTAAAGAAACAC	855
Qy	721	TCAGCTGCAAGCTAGCTACTGGACACATGTTGTTGGCAATGCTGCTGTGATGAGACG	780
Db	856	TCTTCACCTAAATGGCTTCAGGAGCCACACATTTGTCGGAATGTCTATAGTGGATGAATCT	915
Qy	781	GCCAAAGTTGGAGAAGTTGTCTGATGCTCCTGATGTCGCCATTTGGACCTGGGTGTGT	840
Db	916	GCCAAAGTTGGAGAGGTTGTTGATAGGACCAGATGTTGCAATTTGGTTCGTGTGTG	975
Qy	841	GTGGAGACGGCGTGAGGCTTCCCGCTGCACCTGTCATCGCGCGCGTGCCTATCAAGAAG	900
Db	976	ATTGAGTCTGGAGTTAGACTCTCCCGTTGCACCTGTGATCGGAGGAGTCCGCATCAAGAAA	1035
Qy	901	CATGCTTCGACTCAACACAGCATTAATCGCTGGCACTCAACTGTGGTCAATGGGCACGG	960
Db	1036	CATGATGCATCTCAGGTAGCATCATTTGGCTGGCACTCTACTGTTGGACAATGGGCTCGT	1095
Qy	961	ATAGAGAATATGACTATCTCTGGGGAGGATGTTCTGTGTGTGATGAGGTGTACAGCAAT	1020
Db	1096	GTTGAGACATGACCATTCTCGGGGAAGATGCCATGTTTGTGATGAAATTTACAGCAAT	1155
Qy	1021	GCGGTGTGTTTCTCCCAATAAAGAGATCAAGTCAAGCATCTCAAGCCTCAGATCGTC	1080
Db	1156	GGAGGTGTAGTTTCCCCCAACAGGAGATCAATCCAGTATATTGAAACCTGAAATCGTG	1215
Qy	1081	ATGTGA 1086	
Db	1216	ATGTGA 1221	
RESULT 6			
ID	AAC39741 standard; DNA; 556 BP.		
XX	AAC39741;		
AC	AAC39741;		
XX	17-OCT-2000 (first entry)		
DT	Zea mays DNA fragment SEQ ID NO: 25732.		
DE	Zea mays DNA fragment SEQ ID NO: 25732.		
XX	Hybridisation assay; genetic mapping; gene expression control;		
KW	Hybridisation assay; genetic mapping; gene expression control;		

KW	protein identification; signal transduction pathway; metabolic;		
KW	pathway; promoter; termination sequence; corn; ss.		
XX	Zea mays subsp. mays.		
XX	EP1033405-A2.		
XX	06-SEP-2000.		
XX	25-FEB-2000; 2000EP-0301439.		
XX	25-FEB-1999; 99US-0121825.		
PR	05-MAR-1999; 99US-0123380.		
PR	09-MAR-1999; 99US-0123548.		
PR	23-MAR-1999; 99US-0125788.		
PR	25-MAR-1999; 99US-0126264.		
PR	29-MAR-1999; 99US-0126785.		
PR	01-APR-1999; 99US-0127462.		
PR	06-APR-1999; 99US-0128234.		
PR	08-APR-1999; 99US-0128714.		
PR	16-APR-1999; 99US-0129845.		
PR	19-APR-1999; 99US-0130077.		
PR	21-APR-1999; 99US-0130449.		
PR	23-APR-1999; 99US-0130510.		
PR	28-APR-1999; 99US-0131449.		
PR	30-APR-1999; 99US-0132048.		
PR	30-APR-1999; 99US-0132407.		
PR	04-MAY-1999; 99US-0132484.		
PR	05-MAY-1999; 99US-0132485.		
PR	06-MAY-1999; 99US-0132486.		
PR	06-MAY-1999; 99US-0132487.		
PR	07-MAY-1999; 99US-0132863.		
PR	11-MAY-1999; 99US-0134256.		
PR	14-MAY-1999; 99US-0134218.		
PR	14-MAY-1999; 99US-0134219.		
PR	14-MAY-1999; 99US-0134221.		
PR	18-MAY-1999; 99US-0134370.		
PR	19-MAY-1999; 99US-0134941.		
PR	20-MAY-1999; 99US-0135124.		
PR	21-MAY-1999; 99US-0135353.		
PR	24-MAY-1999; 99US-0135629.		
PR	25-MAY-1999; 99US-0136021.		
PR	27-MAY-1999; 99US-0136392.		
PR	28-MAY-1999; 99US-0136782.		
PR	01-JUN-1999; 99US-0137222.		
PR	03-JUN-1999; 99US-0137528.		
PR	04-JUN-1999; 99US-0137502.		
PR	07-JUN-1999; 99US-0137724.		
PR	08-JUN-1999; 99US-0138094.		
PR	10-JUN-1999; 99US-0138540.		
PR	10-JUN-1999; 99US-0138847.		
PR	14-JUN-1999; 99US-0139119.		
PR	16-JUN-1999; 99US-0139452.		
PR	16-JUN-1999; 99US-0139453.		
PR	17-JUN-1999; 99US-0139492.		
PR	18-JUN-1999; 99US-0139454.		
PR	18-JUN-1999; 99US-0139455.		
PR	18-JUN-1999; 99US-0139456.		
PR	18-JUN-1999; 99US-0139457.		
PR	18-JUN-1999; 99US-0139458.		
PR	18-JUN-1999; 99US-0139459.		
PR	18-JUN-1999; 99US-0139460.		
PR	18-JUN-1999; 99US-0139461.		
PR	18-JUN-1999; 99US-0139462.		
PR	18-JUN-1999; 99US-0139463.		
PR	18-JUN-1999; 99US-0139750.		
PR	18-JUN-1999; 99US-0139763.		
PR	21-JUN-1999; 99US-0139817.		
PR	22-JUN-1999; 99US-0139899.		
PR	23-JUN-1999; 99US-0140353.		
PR	23-JUN-1999; 99US-0140354.		

```
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0148341.
PR 09-AUG-1999; 99US-0148565.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149920.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150804.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.

PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155130.
PR 23-SEP-1999; 99US-0155406.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 40.8%; Score 443; DB 21; Length 556;
Best Local Similarity 98.9%; Pred. No. 3.6e-131;
Matches 457; Conservative 0; Mismatches 1; Indels 4; Gaps 1;

QY 1 ATGAAGGCCCTCATCTTGTGCGGGGTTTCGGAACCGCCCTTCGGCCCTTGACTCTGAGC 60
Db 99 ATGAAGGCCCTCATCTTGTGCGGGGTTTCGGAACCGCCCTTCGGCCCTTGACTCTGAGC 158
QY 61 TTCCCGAAACCCCTCGTGGATTTTGCAACAAGCCCATGATTTCTGCACAGATCGAAGCT 120
Db 159 TTCCCGAAACCCCTCGTGGATTTTGCAACAAGCCCATGATTTCTGCACAGATCGAAGCT 218
QY 121 TTGAACAAGCTTGGGGTTCACAGAGCTGGTTTGGCTATCAACTATGCGCCCAAGAGTAATG 180
Db 219 TTGAAGAAGTTGGGGTTCACAGAGGTGGTTTGGCTATCAACTATGCGCCCAAGAGTAATG 278
QY 161 ATTAAATTCCTGAAGGACTTGAGGATAAGCTTGAGGATAAGCTTGAGGATAAGCTTGAGGATAAG 240
Db 279 ATTAAATTCCTGAAGGACTTGAGGATAAGCTTGAGGATAAGCTTGAGGATAAGCTTGAGGATAAG 338
QY 241 ACTGAGCCCTTAGGAACCGCTTGGCCCTTCTTGCTCTAGCAAGGGACAAAGCTTGGCGGATGGA 300
Db 339 ACTGAGCCCTTAGGAACCGCTTGGCCCTTCTTGCTCTAGCAAGGGACAAAGCTTGGCGGATGGA 398
QY 301 TCTGGCCAGCCATCTTTTGTCTCAACAGATGATCTATAGCGAATACCCATTGCTGAA 360
Db 399 TCTG----GCCATCTCTTGTCTCAACAGATGATCTATAGCGAATACCCATTGCTGAA 454
```

QY 361 CTCATCAAAATTCACAAGTGTCTATGGTGGTGAGGCAACAATATATGGTCACTAAGGTGGAT 420  
|||||  
Db 455 CTCATCAAAATTCACAAGTGTCTATGGTGGTGAGGCAACAATATATGGTCACTAAGGTGGAT 514  
|||||  
QY 421 GAACCATCAAAATACGGTGTGTGGTTATGAGGAGGCAACT 462  
|||||  
Db 515 GAACCATCAAAATACGGTGTGTGGTTATGAGGAGGCAACT 556  
|||||  
RESULT 7  
ABL20809  
ID ABL20809 standard; DNA; 1358 BP.  
XX  
AC ABL20809;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 13900.  
DE  
XX Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical; gene; ds.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US09231.  
XX  
PR 23-MAR-2000; 2000US-191637P.  
PR 11-JUL-2000; 2000US-0614150.  
XX  
XX (PEKE ) PE CORP NY.  
XX  
XX Venter JC, Adams M, Li PWD, Myers EW;  
PI  
XX WPI; 2001-656860/75.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
XX  
XX Claim 1: SEQ ID NO 13900; 2lpp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins  
CC (ABB57737-ABB72072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 1358 BP; 375 A; 317 C; 361 G; 305 T; 0 other;  
SQ  
Query Match 38.4%; Score 417.2; DB 23; Length 1358;  
Best Local Similarity 62.8%; Pred. No. 1.1e-122;  
Matches 682; Conservative 0; Mismatches 398; Indels 6; Gaps 2;  
QY 1 ATGAAGGCCCTCATCTGTTCGGGGTTTCGGAACCCGCCCTTCGGCCCTTGTACTCTGAGC 60  
|||||  
Db 279 ACGAGGGCGTGTATCTGTGTCGGTGTATGGACCCCGCTGCGACCCCTGACCTCAGC 338  
|||||  
QY 61 TTCCGGAACCCCTCGTGGATTTTGCACAAAGCCCATGATCTGCACCATCAAGCT 120  
|||||  
Db 339 ACGCCCAAGCAGCGGTGGAGTTTGCACAAAGCCGATCTTCTCCACCACTGGAGGCA 398  
|||||  
QY 121 TTGAAGAAGTTGGGGTACAGAGGTGGTTTTGGCTATCAACTATCGCCAGAGGTAATG 180  
|||||

Db 399 CTCGTGATGCGGGATGTCTGTCAGGTTATTTTAGCCGTACGTATCGCGGACGAATG 458  
|||||  
QY 181 ATTAATTTCTTGAAGACCTTTGAGGATTAAGCTTTGGGATCACAATATACATGCTCCCAAGAG 240  
|||||  
Db 459 GAAAAGGAGCTTAAGTCGAGCAAGAACTGGCGTGGAATGATCTTCTACACGAG 518  
|||||  
QY 241 ACTGAGCCCTTAGGAACCGCTGCCCTCTTGTCTTAGCAAGGACAAAGCTTGGCGATGGA 300  
|||||  
Db 519 ACGGAACCCCTGGGAACAGCTGGACCTCTGGCCCTAGCAAAAACGATTTTAGCAGCCAGC 578  
|||||  
QY 301 TCTGGCCAGGCATCTTTTCTCAACAGTGTATTAAGCCGAATPACCCATTTGCTGAA 360  
|||||  
Db 579 TCAG---AGCCATTTTTCGTGCTCAATTCGGACGTTATATGCGATTTTCCATTTAAACAA 635  
|||||  
QY 361 CTCATCAAAATTCACAAGTGTCTATGGTGGTGAGGCAACAATATATGGTCACTAAGTGGAT 420  
|||||  
Db 636 CTAGTCCAAATTCATTTGTAATCAGGAAGAAGAGGCACAATTTGTTGCACAAGAGTGCAG 695  
|||||  
QY 421 GAACCATCAAAATACGGTGTGTGGTTATGGAGGAGGCAACTGGCAGGGTGGAAAGTTT 480  
|||||  
Db 696 GAACCATCAAAATACGGAGTTGTG---CTGTACGATGAGAAGCGCTGTATAAAAAACTTT 752  
|||||  
QY 481 GTTGAGAAGCCAAAATATTTTGGGTGTAACAAGATCAATGCTGGGATTTACTTACTGAAC 540  
|||||  
Db 753 ATTGAAAAGCCACAGAGTTCTGTAGCAATAAGATCAATGCCGCAATTTACATCTTTAAT 812  
|||||  
QY 541 CCATCTGTCTTACCGCATTTAGCTGAGCGCCCAACATCAATTTGAGAAAGAGTCTTCCCT 600  
|||||  
Db 813 CCGTCGTGCTTGAACGGATTCGAAGTTAAGCCACATCAATAGAGAAGAGGTTGTTCCCT 872  
|||||  
QY 601 CAAATTCAGCTGATCAACAGCTCTATGCAATGCTCTCCAGTGTTCAGAGTGTTCGATGGATGTT 660  
|||||  
Db 873 GAATGACCGCAACAGGAGTTGTATGCAATGGAATTAAGTCTGATGGAGCATC 932  
|||||  
QY 661 GGTACGCTTAGGAGTACATTTACTTGTGCTTGTCTTATCTAGACGATGATAGAGAA 720  
|||||  
Db 933 GGACAGCCAAAAGACTTCCCTAACCGCATGTGTCTGTATCTAAGCTCGCTGCGCCAGAG 992  
|||||  
QY 721 TCAGCTGCCAAGCTAGCTACTGAGCAGCATGTTTGGCAATGCTGTGGTGCATGAGAGC 780  
|||||  
Db 993 CAATCCCCCAAGCTGTACACAGGCCCTGGAGTGGTGGCAACGTTGTTGGGATPCCACG 1052  
|||||  
QY 781 GCCAAGATTGGAGAAGTTGTCTGATTGTCCTGATGTCGCCATTTGGACCTGGGTGTGTT 840  
|||||  
Db 1053 GCCAAATCGCGAGGTTGTCOCATAGGGCCTAATGTAACCATTTGACCGCGCTGCTC 1112  
|||||  
QY 841 GTGGAGCGCGGTGAGGCTTTCCCGCTGCATGTCATGCGCGCGTGCATCAAGAAG 900  
|||||  
Db 1113 ATCGAGGATGGCGTGTGCATTAAGCGCTCGACCATCTCAAAGGCGCATCGTTCGCTCA 1172  
|||||  
QY 901 CATGCTTGATCTCAACAGCATTTATCGGCTGGCAGTCAACTGTTGTCATATGGGACGG 960  
|||||  
Db 1173 CACTCATGCTGGAGTCTGTCATTTGCGTTGCGTGTCTACAGTGGCCGCTGGTTCGC 1232  
|||||  
QY 961 ATAGACAATATGACTACTCTCGGGGAGGATGTTTCATGTTGTGATGAGGTGACAGCAAT 1020  
|||||  
Db 1233 ATCGAAGGATCACCGTGTAGCGGAGGATGTAATCGTGAAGATGAGCTCTACATTAAT 1292  
|||||  
QY 1021 GCGGTGTTGTTCTCCACATAAAGAGATCAAGTCAAGCATTTCTGAAGCCCTGAGATCGTC 1080  
|||||  
Db 1293 GGAGGCCAGTCTTGGCCCCATAAAGCATTTGCGGCCAGTGTCCCAAGAACCCGAGATCAT 1352  
|||||  
QY 1081 ATGTGA 1086  
Db 1353 ATGTGA 1358  
|||||

RESULT 8  
ABL02105  
ID ABL02105 standard; cdna; 1476 BP.  
XX  
AC ABL02105;





cholesterol ester storage; systemic lupus erythematosus; infection; severe combined immunodeficiency; malaria; autoimmune disorder; asthma; allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound; bone damage; cartilage damage; antiinflammatory disease; coagulation; thrombosis; contraceptive; ss.

**Homo sapiens.**

WO200058473-A2.

05-OCT-2000.

31-MAR-2000; 2000WO-US08621.

31-MAR-1999; 99US-0127607.

02-APR-1999; 99US-0127636.

05-APR-1999; 99US-0127728.

30-MAR-2000; 2000US-0540763.

(CURA-) CURAGEN CORP.

Shimkets RA, Leach M;

WPI; 2000-602362/57.

P-PSDB; AAB41757.

Novel nucleic acids and peptides derived from open reading frame X, useful for treating e.g. cancers, proliferative disorders, neurodegenerative disorders and cardiovascular disease.

Claim 5; Page 2268-2269; 5507pp; English.

AAC74446 to AAC77606 encode the proteins given in ABA04237 to ABA43397, which represent the human ORFX open reading frames 1 to 3161. The ORFX sequences have activities such as: cytostatic; hepatotropic; vulnuary; isopteric; antiparkinsonian; nootropic; immunoprotective; osteopathic; anticonvulsant; antiarthritic; neurooppressive; immunostimulant; cardiac; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive; dermatological; immunosuppressive; antiinflammatory; antibacterial; antiviral; antifungal; antineumatic; antithyroid; and antianemic. The sequences can be used for determining the presence of or predisposition to, or preventing or treating pathological conditions associated with an ORFX-associated disorder. The nucleic acids can be used to express ORFX proteins in gene therapy vectors. The proteins and nucleic acids may be used to treat cancers, proliferative disorders, neurodegenerative disorders, osteoarthritis, graft vs host disease, cardiovascular disease, diabetes mellitus, erythematosis, hypothyroidism, cholesterol ester storage, systemic lupus erythematosus, severe combined immunodeficiency (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to enhance coagulation; to inhibit thrombolysis; and as a contraceptive.

Sequence 1512 BP; 323 A; 408 C; 488 G; 292 T; 1 other;

Query Match	36.8%	Score 399.6;	DB 21;
Best Local Similarity	61.1%	Pred. No. 5.2e-117;	Length 1512;
Matches 664;	Conservative	0;	Indels
Mismatches 419;			Gaps 1;

1 ATGAAGGCCCTCATTTCTTGTCGGGGGTTTCGGAACCCGCCCTTCGGCCCTTTGACTCTGAGC 60

142 ATGAAGGCACTGATCTTAGTGGGGGCTATGGGACGGGCTACGGCCGCTGACGCTGAGC 201

61 TTCCCGAAACCCCTCGTGGATTTTGCCAAACAAGCCCATGATTCTGCACCAGATCGAAGCT 120

202 ACCCCGAAGCCACTGGTGGACTTCTGCAATAAGCCCATCTTGCTGCACCAAGTGGAGGCG 261

121 TTGAAAGAACTTGGGGTCACAGAGGTGGTTTGGCTATCAACTATCGCCCGAGAGGTAATG 180

262 CTAGCCGGCGCAGGGCGTGGACCCACGTGATCCTGGCCCGTGAGCTACATGTCGCGAGTGCTG 321

181 ATTAATTTCTTGAAGGACTTTGAGGATAAGCTTGGCATCACAAATTACATGCTCCCAACAC 240

Db	322	GAGAGGAATGAAGGCACAGGACAGAGGCTGGGAATCCGAATCTCCATGTCCCATGAA	381
Qy	241	ACTGAGCCCTTAGGAACCGCTGGCCCTCTTGCTCTAGCAAGGACAAGCTTCGGATGGA	300
Db	382	GAGAGCCTTTGGGACAGCTGGCCCTGGCCCTGACCTACTCTC---TGAG	438
Qy	301	TCTGGCAGCAGATCTTTTGTCTCAACAGTGATGTCAATAGGCAATACCCATTTGCTGAA	360
Db	439	ACTGCAGACCCCTTTCTTCGTCTCAACAGTGAGCTGATCGGATTTCCCTTCCAAGCC	498
Qy	361	CTCATCAAAATTCACAAGTGTGATGGTGTGAGGCAACAATATATGGTCACTAAGGTGAT	420
Db	499	ATGGTGCAGTTTCCACCGGCACCATGGCCAGAGGGCTCCATCCTGGTGACCAAGGTGGAG	558
Qy	421	GAACCATCAAAATACGGTGTGTGGTATGAGGAGGCAACTGGCAGGGGTGCAAAAGGTTT	480
Db	559	GAACCCCTCCAAGTACGGTGTGGTGTGTGAGGCTGACACAGGCGCATTCACCGGTTT	618
Qy	481	GTTGAGAAGCCAAAATATTTTGGGTAAACAAGATCAATGCTGGGATTTACTTACTGAAC	540
Db	619	GTGAGAAAGCCACAGGTGTGTGTGTCCAATAGATCAAGCAGGCATGTACATCCTGAGC	678
Qy	541	CCATCTCTCTTGACCCATTGAGCTGAGGCCAACATCAATTTGAGAAAGAGCTTTCCTT	600
Db	679	CCTCAGTGTGCGGCGCATCCAGCTGCAGCCTACGCTCCATTGAGAAAGAGCTCTCCCC	738
Qy	601	CAAAATGCGAGCTGATCAACAGCTCTATGCAATGGTCCCTTCCAGGTTTTTGGATGATGTT	660
Db	739	ATTTATGCCAAGGAGGGGAGCTATATGCCATGGAGTTACAGGGCTTCGTGATGGACATT	798
Qy	661	GGTCAGCCTAGGGACTACATTACTGGCTTGCCTTTATCTAGACTCGATTAGGAAGAA	720
Db	799	GGGCAAGCCCAAGGACTTCTCACTGGCATGTGCTCTCTCTCGAGTCACGTAGGCGAGAAG	858
Qy	721	TCAGCTGCCAAGCTAGCTACTGGAGCACATGTTGTTGGCAATGTGCTGGTGCATGAGAGC	780
Db	859	CAGCCTGAGCGGCTGTCTCAGGCGCTGGCATTTGTGGGCAACGTGCTGTGAGCCCAAGT	918
Qy	781	GCCAAAGATTGGAGAAGGTTCTCTGATTGGTCTGATGTGCCCATTTGGACCTGGGTGTGTT	840
Db	919	GCCCGCATCGGCCAGAATCGACAGTTGGCCCCAATGTGAGCCTGSGACCTGGCGTGGTG	978
Qy	841	GTGAGAGACGGGTGAGGCTTTCGCCGTGCAGTGTCAATCGGGGGCGTGGGTATCAAGAAG	900
Db	979	GTCGAAGATGGTGTGTATATCCGGCGGTGCACGGTCTCGGGGATGCCCGGATCCGTTC	1038
Qy	901	CATCCTTGCATCTCAACACAGATATACGGCTGGCATCAACTGTTGGTCAATGGGCACGG	960
Db	1039	CATTCTTGGCTTGAGTCTCTGATTTGGCGTGGCGCTGCGCGTGGGTGAGGTACGCTACGC	1098
Qy	961	ATAGAGAATATGACTATCTTGGGGAGGATGTTCAATGTGTGATGAGGTGTACAGCAAT	1020
Db	1099	ATGAGAAAGGTGACAGTGTCTGGGTGAGGACGTCAATAGTTAATGATGAGCTCTACCTCAAC	1158
Qy	1021	GGCGGTGTGTCTCCACATAAAGAGATCAAGTCAAGCATTTCTGAAGCCTCGAGATCGTC	1080
Db	1159	GGAGCCAGCGTGTGCTCCCAACAGTCTATTGGGAGTCAAGTCCAGAGCCTCGTATCATC	1218
Qy	1081	ATGTGA	1086
Db	1219	ATGTGA	1224

RESULT 10	
AAZ94202	
ID	AAZ94202 standard; cDNA; 1559 BP
XX	
XX	
XX	AAZ94202;
XX	
XX	
DT	19-JUN-2000 (first entry)
XX	
DE	Human transferase TRNSFS-2 cDNA





XX Primer sets for synthesizing polynucleotides, particularly the 5602  
PT full-length cDNAs defined in the specification, and for the detection  
PT and/or diagnosis of the abnormality of the proteins encoded by the  
PT full-length cDNAs -  
XX  
XX Claim 8: SEQ ID 18794; 2537bp + CD ROM; English.  
XX  
CC The present invention describes primer sets for synthesizing 5602  
CC full-length cDNAs defined in the specification. Where a primer set  
CC comprises: (a) an oligo-dm primer and an oligonucleotide complementary  
CC to the complementary strand of a polynucleotide which comprises one of  
CC the 5602 nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in  
CC the specification. The primer sets can be used in antisense therapy and  
CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.  
XX  
XX Sequence 3220 BP: 631 A; 865 C; 1035 G; 689 T; 0 other;

Query Match 32.9%; Score 357.4; DB 22; Length 3220;  
Best Local Similarity 61.5%; Pred. NO. 2.7e-103;  
Matches 590; Conservative 0; Mismatches 366; Indels 3; Gaps 1;

Qy 1 ATGAGCGCCCTCATTTCTGCGGGCTTTCCGGAACCCGCTTCGGCTTTGACTCTGAGC 60  
Dd 226 ATGAGGCACTGATCTAGTGGGGGCTATGGAGCGGCTAGCGCGCTGACGCTGAGC 285  
Qy 61 TCCCGAAACCCCTCGTGGATTTGGAACAGCCCATGATTTGACCAAGATCGAAGCT 120  
Dd 286 ACCCGCAAGCCACTGTGAGCTTCTCAATAGCCCATCTTGTGACCAAGTGGAGGCG 345  
Qy 121 TTGAGAGAGTTGGGTGACAGAGTGGTTTGGCTATCAACTATGCCCGAGAGTAATC 180  
Dd 346 CTAGCGCGCGCGCGCTGACACACCTGATCCTGGCGCTGAGCTACATGTGCGAGGTGCTG 405  
Qy 181 ATTAATTTCTTGAAGGACTTTTGAGGATAAGCTTTGGCATCACAAATTACATGCTCCGAAGAG 240  
Dd 406 GAGAGGAATGAGGCACAGGACAGAGGCTGGGAAATCGGAATCTCCATGTCCCATGAA 465  
Qy 241 ACTGAGCCCTTAGGAACCCCTGCGCTCTGCTCTACCAAGGACAGGTTGGCGGATGCA 300  
Dd 466 CAGGACCCCTTTGGGACACAGCTGGCCCTCGGCTGCGCGCTGACCTACTCTC--TGAG 522  
Qy 301 TTGCGCCAGCCATCTTTGCTTCAACAGTGTGTGATAGCGAATACCCATTTGCTGAA 360  
Dd 523 ACTGCGAGACCCCTTCTCTGCTCTCAACAGTGTGTGATAGCGAATTTGCTGAA 582  
Qy 361 CACATCAAAATTTTCAACATCTCATGCTGTTGAGGACACAAATTTGCTACTTAGGTGGAT 420  
Dd 583 ATGTTGACATTTTCCACGGGACAGATGCGGAGAGGGCTCCATCTGTGACCAAGAGTGGAG 642  
Qy 421 GAACATCAAAATACCGCTGTGTGCTTTAGGAGGACGCAAGCTGCGAGGTTGCAAGCTTT 480  
Dd 643 GAACCTTCCAAAGTACGGTGTGTGTTGTTGTGTGAGGCTTGTGACAGGCGCGCATTCACCGGTT 702  
Qy 481 GTTGTGAGGCAAAATATTTTGGGTAAACAAGATCAATGCTGGGGATTTACTTTACATCAAC 540  
Dd 703 CTGCGAAGAGCCACAGCTGTTTCTGCTCAATAGATCAACCGCAGGCAATGTACATCTCTGAGC 762

Qy 541 CCATCTCTCTCTGAGCGGCTGAGCTGAGCGGCAACATCAATTCAGAAAGAGTCTTCCCT 600  
Dd 763 CCTCAGCTGTCTGCGCGCGCATCCAGCTGAGCGCTAGCTGCTCCATTTGAGAGAGAGTCTTCCCC 822  
Qy 601 CAAATTGCAAGCTGATCAACAGCTCTATCAATGCTCTTCCAGGCTTTTGGATGCAATGTT 660  
Dd 823 ATTTATGCGCAAGAGGGGCGAGCTATATGCCATGGAGTTACAGGGCTTCTGGATGACAT 882  
Qy 661 GGTGAGCTTAGGGACTACATTTACTTGGCTTGGCTTTTATCTAGACTCGATTAGSAGAAA 720  
Dd 883 GGGCAGCCCAAGGAGCTTCTCCTCAGTGGCATGCTCTTCTTCTCAGTCACTGAGGCGAGAAG 942  
Qy 721 TCAGCTGCGCAAGCTAGCTTACTTGGAGCACATGTTGTTGCAATGCTGCTGGTGCATGAGAGC 780  
Dd 943 CAGCTGAGCGGCTGTGTCTAGGCGCTTGGCATTTGTTGGGCAACGTTGCTGACACCAAGT 1002  
Qy 781 GCCAAGATGGAAGGCTGTCTGATGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
Dd 1003 GCGCGCATGCGGCCAAGACTGCGAGCATTTGGCCCCCAATGTGAGCCTGGGAGCTTGGCGTGGTG 1062  
Qy 841 GTGAGGAGCGGCTGAGGCTTTCCGCTGCACTGTCTGATGCGCGGCTGCTGCTGCTGCTGCT 900  
Dd 1063 GTCGAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1122  
Qy 901 CATGCTTGCATTCACAAAGCAATTTATCGGCTGCGCACTCACTGTTGGTCAATGGGACAG 959  
Dd 1123 CATCTCTGCTTGTAGTCTTGCATTTGGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 1181  
RESULT 13  
AAS41146  
ID AAS41146 standard; cDNA; 990 BP.  
XX  
AC AAS41146;  
XX  
DT 17-DEC-2001 (first entry)  
XX  
XX cDNA encoding novel human enzyme polypeptide #362.  
XX  
KW Human: oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;  
KW ligase; hyperproliferative disorder; immunodeficiency disorder;  
KW autoimmune disorder; neurological disorder; metabolic disorder;  
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;  
KW blood-related disorder; infectious disorder; gene therapy; cytostatic;  
KW anti arthritic; nephrotropic; anticoagulant; ss.  
XX  
OS Homo sapiens.  
XX  
DN WO200155301-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001W0-US01239.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 11-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.



CC AAS40785-AAS41684 represent cDNA sequences encoding for the novel human  
CC enzyme polypeptides of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX  
SQ Sequence 990 BP; 188 A; 272 C; 325 G; 196 T; 9 other;  
  
Query Match 29.7%; Score 322; DB 22; Length 990;  
Best Local Similarity 61.5%; Pred. No. 2.9e-92;  
Matches 598; Conservative 3; Mismatches 342; Indels 5; Gaps 3;  
  
QY 1 ATGAGAGCCCTATTTCTGTCGGGGCTTCGGGAACCCGCTTCGGCTTTGACTCTCGAGC 60  
DB 65 ATGAGGCACTGATCTTAGTGGGGGCTATGGGACGCGGTACGGCGCTGAGCGTGGAGC 124  
QY 61 TTCGGAACCCCTCGTGATTTGCAACAAGCCCATGATTCGACCAAGATCGAAGCT 120  
DB 125 ACCCGCAAGCCATGCTGGCACTTCGCAATGACCCATCTTCTGCACCAAGTGGAGGCG 184  
QY 121 TTGAAGAAGATTGGGTACAGAGGTGGTTTGGCTATCAACTATGCGCCGAGAGGTAATG 180  
DB 185 CT-ANGCGGAGCGCTGGACCACTGATCCTGCGGTGAGCTACATCTCGAGGCTCTG 243  
QY 181 ATTAATTTCTTAAGACATTTGAGTAAGCTTTGGCATCACAAATTACATGCTCCCAAGAG 240  
DB 244 GAGAAGGAATGAAGCACAGGAGAGAGGCTGGGAATCGGAATCCGAATCTCCATGTCCCATGAA 303  
QY 241 ACTGAGCCCTTAGCAACCCGCTGCCCTCTTCCTACAGGACGACAAAGCTTGGGGATGGA 300  
DB 304 GAGGAGCCCTTTGGGACAGCTGGGCCCTCGGCGCTGGCGGTGACCTACTCTC---TGAG 360  
QY 301 TCTGGCCAGCCATCTTTGTCTCAACAGTGTGTCATGAAGCAATACCCATTTGCTGAA 360  
DB 361 ACTGCAGACCCCTTCTCTGCTCTCAACAGTGTGTCATGTCGATTCGCTTCCCTCCAAAGC 420  
QY 361 CTCATCAAAATTCACAAGTGTCTAGTGGTGGAGGCAACAATATTGGTCACTAAGGTGGAT 420  
DB 421 ATGGTGCAGTTCCACCGGAGGAGTGGCCAGGAGGGCTCCATCTGCTGACCAAGGTGAG 480  
QY 421 GAACCATCAAAATACCGTGTGTGGTGTATGGAGAGGCAACATCCGCGCTGCGAAAGCTTTT 480  
DB 481 GAACCCCTCAATPACGGTGTGGTGTGTGGAGGCTGACACAGGCGCGCATTCACCGGTTT 540  
QY 481 GTTGAGAAGCAAAATATTGGGTGTAAGATCAAGATCAATGCTGGATTTACTTACTTGAAC 540  
DB 541 GTGGAAGCCACAGGT 600  
QY 541 CCATCTGTCTTTCACCGCATTCAGCTGAGGCCAACATCAATTTGAGAAGAGGTCTTCCCT 600  
DB 601 CCTGCACTGCTGGCGGCATCCAGCTGCAGCTACGCTTCCATTCAGAAGAGGCTTCCCC 660  
QY 601 CAATTCAGCTGATCAACAGCTGTATGCAATGGTCTTTCAGGTGTTTGGATGGATGTT 660  
DB 661 ATTATGCCCAAGGAGGGGCAANNCTATGCCATGGAGTTACAGGGCTTCTGGATGGAGATT 720  
QY 661 GGTGACCTTAGGACTACATTTACTGCTTGGCTTGTATGAGAGGCAATGCTTGGTGCATGAGAG 720  
DB 721 GGGGACCCCAAGGAGTCTCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780  
QY 721 TCAGCTGCCAAGCTACTTACTGAGCAGACATTTGTTTGGCAATGCTTGGTGCATGAGAGC 780  
DB 781 CACCTGAGCGGT-CTGCTGACCCCTGGCATTTGTGGGCAACCTGCTGCTGAGCCCAAGT 839  
QY 781 GCGAAGATTGGAGAGGTTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
DB 840 GCCCGCATCGCCAGAACTGCAGCATTTGGCCCAATGTGAGCCCTGGACCTGCGCGGTG 899  
QY 841 GTGAGGAGCGGTGAGGCTTTCCCGCTGACATGCTCATGCGCGGCTGCTGCTATCAAGAAG 900  
DB 900 CTGCAAGATTCCTCTCTCTATCCGGCGGTGACGCTGCTGCGGATGCTCCGGATCCGTTCC 959  
QY 901 CATGCTTG 908

DB 960 CATTCCTG 967  
RESULT 14  
AAF13173  
ID AAF13173 standard; cDNA; 1351 BP.  
XX  
AC AAF13173;  
XX  
DT 13-MAR-2001 (first entry)  
XX  
DE Aspergillus oryzae EST SEQ ID NO:5696.  
XX  
KW Multiple gene expression; filamentous fungal cell; EST;  
KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;  
KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;  
KW culture condition; environmental stress; spore morphogenesis;  
KW metabolic pathway engineering; catabolic pathway engineering; ss.  
XX  
OS Aspergillus oryzae.  
XX  
PN WO200056762-AZ.  
XX  
PD 28-SEP-2000.  
XX  
PF 22-MAR-2000; 2000WO-US07781.  
XX  
PR 22-MAR-1999; 99US-0273623.  
XX  
PA (NOVO ) NOVO NORDISK BIOTECH INC.  
PA (NOVO ) NOVO NORDISK AS,  
XX  
PI Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;  
XX  
DR WPI; 2000-594572/56.  
XX  
PT Monitoring differential expression of genes in filamentous fungal cells  
PT uses fluorescence-labeled nucleic acids isolated from the cells and a  
PT substrate of expressed sequence tags -  
XX  
PS Claim 88; Page 2361; 3161pp; English.  
XX  
CC The present invention describes a method for monitoring differential  
CC expression of genes in a first filamentous fungal (FF) cell relative to  
CC expression of the same genes in one or more second filamentous fungal  
CC cells. The method uses fluorescence-labeled nucleic acids isolated from  
CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs  
CC are used in the methods for monitoring differential expression of genes  
CC in a first filamentous fungal (FF) cell relative to expression of the  
CC same genes in one or more second filamentous fungal cells. Monitoring  
CC the global expression of genes from FF cells allows the production  
CC potential of the microorganisms to be improved. New genes may be  
CC discovered, possible functions of unknown open reading frames can be  
CC identified and gene copy number variation and stability can be  
CC monitored. The expression of genes can be used to study how FF cells  
CC adapt to changes in culture conditions, environmental stress, spore  
CC morphogenesis, recombination, metabolic or catabolic pathway  
CC engineering. Using ESTs provides several advantages over genomic or  
CC random cDNA clones including elimination of redundancy as one spot on an  
CC array equals one gene or open reading frame, and organization of the  
CC microarrays based on function of the gene products to facilitate  
CC analysis of the results. AAF07478 to AAF1247 represents ESTs from  
CC Fusarium venenatum; AAF1248 to AAF1853 represents ESTs from Aspergillus  
CC niger; AAF1854 to AAF14878 represents ESTs from Aspergillus oryzae; and  
CC AAF14879 to AAF15337 represents ESTs from Trichoderma reesei, which are  
CC all specifically claimed in the present invention.  
XX  
SQ Sequence 1351 BP; 305 A; 365 C; 315 G; 365 T; 1 other;  
  
Query Match 27.9%; Score 300.4; DB 21; Length 1351;  
Best Local Similarity 57.5%; Pred. No. 2.9e-85;  
Matches 604; Conservative 0; Mismatches 431; Indels 15; Gaps 3;

```
QY 49 TTGACTCTGAGCTTCCCGAAGCCCTCGTGGATTTTGCAGAAAGCCCAAGATTTCTGCAC 108
Db 9 TAGACATTGACCTTCCCAAGCCCTGCTGGAAATTTGGGAACCGTCCCATGATTTTGCAC 68
QY 109 CAGATCGAAGCTTTGAAGAAAGTTGGGTCACAGAGGTGGTTTGGCTATCAACTATCGC 168
Db 69 CAAGTCGAGAGCTTGGCTGCTGGTGTACGGGATTTTGTCTGGCTGTTAACTACCGC 128
QY 169 CCAGAGGTAATGATTAATTTCTTTGAAGGACTTTTGAAGATAAGCTTGGCATCACAATTACA 228
Db 129 CCAGATGTCATGTTTTCGCCCTCAAGAGATATGAGGAACAATGTGAATCGAG 188
QY 229 TGCTCCCAAGAGATGAGCCCTTAGGAACCGCTGGCCCTCTTCTCTAGCAAGGGACAAG 288
Db 189 TTTCTCCGTCCGATCCGAACCCCTGGGCACTGCGCGTCCCTGAAAGTTAGC--GGAAAAG 245
QY 289 CTTGGCGATGATCTGGCCAGCCATCTTTGTCTTCAACAGATGATGTATAGCGAATAC 348
Db 246 ATTTTGGGCAAGGACGACTCTCCATTTCTGTTCTTAACTCCGATGTCATCTGCGACTAC 305
QY 349 CCATTGTCTGAACATCAATAATTTCCAAAGTGTCTATGTTGGTGAGCAACAATTTATGGTC 408
Db 306 CCTTCAAGGACTGGCTGAGTTCCATAGAAGCATGGCGATGAAGGTACCATTTGTGTT 365
QY 409 ACTAAGTGGATGAACCATCAAAATACGGTGTGTGGTT--ATGGAGGAGCAACTGGC 465
Db 366 ACTAAGTTCGAGGACGCCCTCCAAGTATGCGGTGTTGTTCACAAGCCCAATCAACCCCTCG 425
QY 466 AGGTGGAAAGTTTGTGTAGAGGCCAATAATTTTGGTGGGTAAAGATCAATGCTGGG 525
Db 426 CGCATCGACCGCTTGTGCGAAAAGCTGTTGAATTCGTTGGAACCGCATCAACGCCGCT 485
QY 526 ATTTACTTACTGAACCCATCTCTCCCTGACCGCATGAGCTGAGGCCCAACATCAATTGAG 585
Db 486 ATCTACATCATGAACCCCTAGTGTGTGAATCGCATGAGCTGCGCCGACATCCCATTTGAG 545
QY 586 AAAGAGGCTTCCCTCAAAATGTCAGCTGATCAACAGCTCTATGCAATGGTCTCTCCAGGT 645
Db 546 CAGGAGACATTCGCGCTCTCTGCAAGGACGACAGCTCCATCTCTTCACTGGAGGT 605
QY 646 TTTGGATGGATTTGGTCAAGCTAGGACTACATTTACTGGGTGGTCTTTATCTAGAC 705
Db 606 TTTGGATGGAGTTGGTCAACCTAGGACTTCTCAGCGGCACTGTCTATACCTCACT 665
QY 706 TCGATTAGGAAGAAATCAGCTCCCAAGCTAGCTACTGGAGCACATGTT-----GTT 756
Db 666 TCTCTCGCAAGCGCAATTCGAAGCTGCTCGCACCCCACTCCGAACCCCTACGTGTACGGT 725
QY 757 GCAATGTGCTGGTGCATGAGAGCGCCCAAGATTGGAGAAGTTGTCTGATTTGGTCTGTAT 816
Db 726 GCAATGTTATGTTGACCCCTCCGCCAAGATTGCAAGAATTGGCGCATTTGGTCCCAAT 785
QY 817 GTCGCCATTGACCTGGTGTGTGTGGAGGACGCGTGTAGGCTTTCCCGCTGCACTGTC 876
Db 786 GTAGTCATTGTTCCCAACGTTGTGTGCTGATGAGTTCTGCTGAGGTTGTGTCGCTT 845
QY 877 ATGCGCGGCTGCTGATCAAGAAAGATGCTTGCATCTCAAGCAGATTTATCGGTGGCAC 936
Db 846 CTAGAGAACAGCAAGGTCAAGACCATGCTGGGTTAAGTCAACCATTTGTCGGATCGAAT 905
QY 937 TCAACTGTTGTTCATGGCAGGATAGAGATATGACTATCTCTGGGAGGATGTTTAT 996
Db 906 AGCTCCGTAGGAGGTGGGCTGTTTGGAGAATGTACCGTCTGGGTGACGACGTTACT 965
QY 997 GTGTGTGATGAGGTACAGCAATGGCGTGTGTTCTCCCAACATAAAGAGATCAAGTCA 1056
Db 966 ATCGACAGCAAGTCTATGTTAATGGTGGATCCATCCCTCTCATAGAGCATCAAGCAG 1025
QY 1057 AGCATTTGAAGCCCTGAGATCGTCAATGTA 1086
Db 1026 AACGTTGATGTTCCCGCATATTATTATGTA 1055
```

```
RESULT 15
AAC43567
ID AAC43567 standard; DNA; 458 BP.
XX
AC AAC43567;
XX
DT 17-OCT-2000 (first entry)
XX
DE Zea mays DNA fragment SEQ ID NO: 39699.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic;
KW pathway; promoter; termination sequence; corn; ss.
XX
OS Zea mays subsp. mays.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
```



```
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146380.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.

PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149923.
PR 21-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 24.3%; Score 262.6; DB 21; Length 458;
Best Local Similarity 96.8%; Pred. No. 2e-73;
Matches 268; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ATGAAGCCCTCATTTCTCGGGGGTTTCGGAACCCCGCTTCGGCCCTTTACATCTGACC 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 182 ATGAAGCCGCTCATTTCTCGGGGGTTTCGGAACCCCGCTTCGGCCCTTTACATCTTAGC 241
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 TTCCCGAAACCCCTCGTGGATTTTGCAACAAAGCCCATGATTCGCACCATCGAAGCT 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 242 TTCCCGAAAGCCCTTCTGATTTTCCAAAACAGCCCATGATTCGCATCATCGAAGCT 301
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 TTCAAAAGACTTGGGGTTCACAGAGGTGGTTTGGCTATCAACTATCGCCAGAGTAATG 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 302 TTGAAGAGATTGGGGTTCACAGAGGTGGTTTGGCCATCACTACCGCCAGAGTAATG 361
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: November 27, 2002, 02:46:27  
Job time : 264 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
C	1	74.2	6.8	1535	3	US-09-032-372-8	Sequence 8, Appl
	2	73.2	6.7	787	4	US-09-007-119-20	Sequence 20, Appl
	3	58.4	5.4	7902	4	US-08-961-527-112	Sequence 112, App
	4	55.2	5.1	690	3	US-09-024-023-1	Sequence 1, Appl
	5	55.2	5.1	690	3	US-09-531-111-1	Sequence 1, Appl
	6	53.6	4.9	287	4	US-09-034-023-3	Sequence 3, Appl
	7	53.6	4.9	287	4	US-09-531-111-3	Sequence 3, Appl
	8	52.8	4.9	2634	1	US-08-196-218-31	Sequence 31, Appl
	9	52.8	4.9	2634	1	US-08-681-953-31	Sequence 31, Appl
	10	46.2	4.3	678	4	US-03-370-838-155	Sequence 155, App
11	41.6	3.8	6854	4	US-03-194-905-7	Sequence 7, Appl	
12	40.4	3.7	3571	4	US-09-221-017B-904	Sequence 904, App	
13	40.2	3.7	28804	2	US-08-592-874-1	Sequence 1, Appl	
14	40.2	3.7	28804	3	US-09-096-942-2	Sequence 2, Appl	
15	40.2	3.7	28804	3	US-09-096-867-2	Sequence 2, Appl	
16	39	3.6	1272	3	US-03-036-987A-28	Sequence 28, Appl	
17	39	3.6	1272	4	US-03-370-700-28	Sequence 28, Appl	
18	37.4	3.4	5144	1	US-08-487-890A-105	Sequence 105, App	
19	37.4	3.4	5144	2	US-08-478-435-105	Sequence 105, App	
20	37.4	3.4	5144	2	US-08-337-483-105	Sequence 105, App	
21	37.4	3.4	5144	2	US-08-478-373-105	Sequence 105, App	
22	37.4	3.4	5144	3	US-08-474-671-105	Sequence 105, App	
23	37.4	3.4	5144	3	US-08-483-577A-105	Sequence 105, App	
24	37.4	3.4	5144	4	US-08-897-438-105	Sequence 105, App	
25	37.4	3.4	5144	4	US-08-637-654-105	Sequence 105, App	
26	37.4	3.4	5144	4	US-08-649-518-105	Sequence 105, App	
27	36.8	3.4	4403765	4	US-09-103-840A-2	Sequence 2, Appl	

```
Query Match          6.8%; Score 74.2; DB 3; Length 1535;
Best Local Similarity 49.5%; Pred. NO. 4e-14;
Matches 255; Conservative 0; Mismatches 248; Indels 12; Gaps 2;

QY 31 GGAACCGCGCTTCGGCTTGTGACTCTGAGCTTCCCGAAACCCCTCGTGGATTTTGCAAC 90
DB 113 GGAACCTCGCTTCAGACCTTGTCTTTGAGGTGCCCAACCATTTGCTCGTGGCAGGG 172

QY 91 AAGCCCATGATTTCTGACACGATCGAAGCTTTGAAAGAAGTTGGGGTCAAGAGGTGGTT 150
DB 173 GTCCCTATGATCAACACCATATTGAGCGCTGTGCCAGGTCCCTGGAATGCGAGAGATT 232

QY 151 TTGGCTATCAACATCGCCCA-----GAGGTAATGATTAATTTCTTGAAGGACTTTGAC 204
DB 233 CTGCTCATTTGGCTTCTACCAACCTGATGAGCCCTTCAACCCAGTTCTTAGAAGCGGCCAG 292

QY 205 GATAAGCTTGGCATCAGATATGATGCTGCCAAGAGACTGAGCCCTTAGAAGCGGTGGC 264
DB 293 CAGGAGCTTTAAACCTTCCACTCAGTACCTGCGAGGAATTTGCCCTTAGGCACAGGGGT 352

QY 265 CTTCTTGGCTCTAGCAGGAGCAAGCTTGGGATGGATCTGGCCAGCCATTCTTTGTCCCTC 324
DB 353 GGTCTTTAGCATTTTCGAGACGACATCTCGCTGGGAGGCCCGCAGGCATTTCTCGCTC 412

QY 325 AAACGCTGATGCTATAAGCGCAATACCAATTTGCTGAACTCATCAAAATTTCAAAAGTGTCT 384
DB 413 AATGCTGATGCTGCTCCGACTTCCCTTTGAGTGCTATGTTGGAAGCCACCGAGCGCAG 472

QY 385 GGTGGTGAAGCAACATTTATGTCACATAAG-----TGGATCAACGATCAAAATACGCT 438
DB 473 CTTCAACCTTTCTTACTCTCTGCACTAGCGCTAACAGCAGCAATCTCTCAACTACGGC 532

QY 439 GTTGTGTTATGAGGAGGCACTGCGAGGTGGAGAGTTTCTTGAGAGCCCAAAATA 498
DB 533 TGCATCGTTGAGAAATCCAGACACAGAGGTATTGCACTATGTGAGAAACCCAGACACA 592

QY 499 TTGTGGCTAAACAGATCAATGCTGGGATTTACTT 533
DB 593 TTATCAGTGACATCACTACCTGCGGCATCTACCT 627

RESULT 2
US-09-007-119-20/c
; Sequence 20, Application US/09007119C
; Patent No. 6300541
; GENERAL INFORMATION:
; APPLICANT: Lightfoot, David A.
; APPLICANT: Gibson, Paul T.
; APPLICANT: Merkem, Khalid
; TITLE OF INVENTION: Soybean Sudden Death Syndrome Resistant Soybeans,
; TITLE OF INVENTION: Soybean Cyst Nematode Resistant Soybeans and Methods of
; TITLE OF INVENTION: Breeding and Identifying Resistant Plants
; FILE REFERENCE: Sou Illinois 1268/2 Sequence Listing
; Patent No. 6300541
; CURRENT APPLICATION NUMBER: US/09/007,119C
; CURRENT FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,335
; EARLIER FILING DATE: 1997-01-14
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 20
; LENGTH: 787
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3)
; OTHER INFORMATION: (a or c or g or t/u)
```

```
; NAME/KEY: misc_feature
; LOCATION: (400)..(401)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (421)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (428)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (432)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (462)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (475)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (514)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (521)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (526)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (532)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (542)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (572)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (574)..(575)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (589)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (603)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (609)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (614)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (636)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
```

```
; LOCATION: (638)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (671)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (687)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (696)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (703)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (708)
; OTHER INFORMATION: (a or c or g or t/u)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (712)
```

Query Match 6.7%; Score 73.2; DB 4; Length 787;

Best Local Similarity 56.0%; Pred. No. 5.7e-14;

Matches 227; Conservative 0; Mismatches 151; Indels 27; Gaps 5;

```
QY 524 GGATTACTTACTGAACCCATCTGCTGACCGCATTTGAGCTGAGCGCCAAACATCAATTG 583
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 563 GGGTAACTGGATGAACCCCTTNGGTATGGANAGATNGAGNGGACACCNATTTCTTGA 504
QY 584 AGAAAGAGTCTTCCTCAAAATGTCAGCTGATCAACAGCTCTATGCAATGGTCTTCC-- 641
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 503 AGACAGAGGGTTTCCAAATATTTTCNGATAGAAAGCTGNATGCAATGGTCTTCCAC 444
QY 642 ----AGGTTTTGGATGGATGTGTGTCAGCCTAGGAGCTACATTAATGCTGCGTCTTT 697
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 443 AGAAGGATTTNAGANGAAATNGGACAA---AGGGACTATATTTNNGANTGAGACTCN 387
QY 698 ATCTAGACTCGATTTAGGAAGA---AATCAGCTGCCAGCTAGTACTTGGAGCACATGTTG 754
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 386 ACNTGAACTGGGAGGAATAATTTCTTAGTGAACNGCCCAATCACCNTCANACANTG 327
QY 755 TTGGCAATGTGTGTGTCATGAGAGCGCCAAAGATTGGAGAAGTTGTCTGATTGGTCCG 814
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 326 TGGAAATGTGTGTGTCATGAGACTGCCAA-ATNGGTGATGATGTCAAATTTGGTCC-- 270
QY 815 ATCTGCCCATTTGACCTGGTGTGTGTGGAGGACCGGTGAGGCTTTCCTCGCTGCACTG 874
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 269 -----ATCCNGGATGTATCATTTGAGAAAGAGTTGGGCTCAAAATCTNACACTG 222
QY 875 TCATGCGCGCGTGCCTATCAAGAAGCATGCTTGCATCTCAAAACA 919
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 221 TTATGTGAAGCACTCGTGTATAAAGCAATTTGTTATATCNGACA 177
```

RESULT 3

US-08-961-527-112

; Sequence 112, Application US/08961527

; Patent No. 6420135

; GENERAL INFORMATION:

; APPLICANT: Charles Kunsch

; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences

; NUMBER OF SEQUENCES: 391

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: Maryland

; COUNTRY: USA

;/ ZIP: 20850  
;/ COMPUTER READABLE FORM:  
;/ MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage  
;/ COMPUTER: HP Vectra 486/33  
;/ OPERATING SYSTEM: MSDOS version 6.2  
;/ SOFTWARE: ASCII TEXT  
;/ CURRENT APPLICATION DATA:  
;/ APPLICATION NUMBER: US/08/961,527  
;/ FILING DATE:  
;/ CLASSIFICATION: 424  
;/ PRIOR APPLICATION DATA:  
;/ APPLICATION NUMBER:  
;/ FILING DATE:  
;/ ATTORNEY/AGENT INFORMATION:  
;/ NAME: BROOKES, A. Anders  
;/ REGISTRATION NUMBER: 36,373  
;/ REFERENCE/DOCKET NUMBER: PB340P1  
;/ TELECOMMUNICATION INFORMATION:  
;/ TELEPHONE: (301) 309-8504  
;/ TELEFAX: (301) 309-8512  
;/ INFORMATION FOR SEQ ID NO: 112:  
;/ SEQUENCE CHARACTERISTICS:  
;/ LENGTH: 7902 base pairs  
;/ TYPE: nucleic acid  
;/ STRANDEDNESS: double  
;/ TOPOLOGY: linear  
;/ US-08-961-527-112

Query Match 5.4%; Score 58.4; DB 4; Length 7902;  
Best Local Similarity 52.5%; Pred. No. 1.6e-08;  
Matches 128; Conservative 0; Mismatches 116; Indels 0; Gaps 0;  
QY 2 TGAAGGCCCTCATCTTGTGCGGGGTTTCGGAACCGGCTTCGGCCTTTGACTCTGAGCT 61  
Db 4961 TGAAGCCATATCTATGACAGCGGATTCGGAACCTGCTGCTATGACTGAAAATA 5020  
QY 62 TCCCGAAACCCCTCGTGGATTTTGCAACAAGCCCCATGATTTCTGCACAGATCGAAGCTT 121  
Db 5021 CCCCTAAAGCCTTGGTTGAGTTAATCAAAAACCTTTGATTGAGTACCAAAATTCAGCTTC 5080  
QY 122 TCAAGAACTTGGGTCACAGAGCTGCTTTTGGCTATCACTATCCGCCAGAGGTAATGA 181  
Db 5081 TCAAGAAAAGGAATCAATGACATCATCATTTTGGTTATCTTTAAGAGACAATTCG 5140  
QY 182 TTAATTTCTTGAAGGACTTTGAGGATAAGCTTGGCATCACAATTCATCATGCTCCCAAGAGA 241  
Db 5141 ATTACTTGAAGAGAAATFACGGTGTTCGCTCTGCTTTTCAATGATAAATACGCTGACTACA 5200  
QY 242 CTGA 245  
Db 5201 ATAA 5204

RESULT 4  
US-09-024-023-1  
; Sequence 1, Application US/09024023  
; Patent No. 6110899  
; GENERAL INFORMATION:  
; APPLICANT: Lonetto, Michael A.  
; TITLE OF INVENTION: licc  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dechert, Price & Rhoads  
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103-2793  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0

;/ CURRENT APPLICATION DATA:  
;/ APPLICATION NUMBER: US/09/024,023  
;/ FILING DATE:  
;/ CLASSIFICATION:  
;/ PRIOR APPLICATION DATA:  
;/ APPLICATION NUMBER: 60/039,210  
;/ FILING DATE: 28-FEB-1997  
;/ ATTORNEY/AGENT INFORMATION:  
;/ NAME: Falk, Stephen T  
;/ REGISTRATION NUMBER: 36,795  
;/ REFERENCE/DOCKET NUMBER: GM50018  
;/ TELECOMMUNICATION INFORMATION:  
;/ TELEPHONE: 215-994-2488  
;/ TELEFAX: 215-994-2222  
;/ TELEX:  
;/ INFORMATION FOR SEQ ID NO: 1:  
;/ SEQUENCE CHARACTERISTICS:  
;/ LENGTH: 690 base pairs  
;/ TYPE: nucleic acid  
;/ STRANDEDNESS: double  
;/ TOPOLOGY: linear  
;/ US-09-024-023-1

Query Match 5.1%; Score 55.2; DB 3; Length 690;  
Best Local Similarity 51.0%; Pred. No. 4.3e-08;  
Matches 126; Conservative 0; Mismatches 116; Indels 0; Gaps 0;  
QY 2 TGAAGGCCCTCATCTTGTGCGGGGTTTCGGAACCGGCTTCGGCCTTTGACTCTGAGCT 61  
Db 2 TGAAGCCATATCTATGACAGCGGATTCGGAACCTGCTGCTATGACTGAAAATA 61  
QY 62 TCCCGAAACCCCTCGTGGATTTTGCAACAAGCCCCATGATTTCTGCACAGATCGAAGCTT 121  
Db 62 CCCCTAAAGCCTTGGTTGAGTTAATCAAAAACCTTTGATTGAATACCAAAATTCAGCTTC 121  
QY 122 TGAAGAACTTGGGTCACAGAGGTTGGTTTGGCTATCACTATTCGCCAGAGGTAATGA 181  
Db 122 TCAAGAAAAGGAATCAATGACATCATCATCTTGGTTATCTTTAAAGACAATTCG 181  
QY 182 TTAATTTCTTGAAGGACTTTGAGGATAAGCTTGGCATCACAATTCATCATGCTCCCAAGAGA 241  
Db 182 ATTATTTAAAGAAAATACGGTGTTCGCTCTGCTTTTCAATGATAAATACGCTGACTACA 241  
QY 242 CTGA 245  
Db 242 ATAA 245

RESULT 5  
US-09-531-111-1  
; Sequence 1, Application US/09531111  
; Patent No. 6303571  
; GENERAL INFORMATION:  
; APPLICANT: Lonetto, Michael A.  
; TITLE OF INVENTION: licc  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dechert, Price & Rhoads  
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103-2793  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/531,111  
; FILING DATE: 17-Mar-2000  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:





```
/
/
/ TELEX: <Unknown>
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 287 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-531-111-3

Query Match          4.9%; Score 53.6; DB 4; Length 287;
Best Local Similarity 51.2%; Pred. No. 8.5e-08;
Matches 125; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

QY  2  TGAAGCCCTCATCTTGTGCGGGGTTTCGCAACGCCCTTCGCGCTTTCACCTCTGAGCT 61
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  2  TGAAGCCATATCTTAGCAGCGGATGGGAACCTCGCTTGCCTATACACTGAAATA 61
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  62  TCCGAAACCCCTCGTGGATTTTGCACCAAGCCCATGATTCGCACAGATCGAAGCTT 121
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  62  CCCTAAAGCCTTGGTTACAGTTAATCAAAACCTTTGATGAATACCAAAATTCGATTC 121
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  122 TGAAGAAGTTGGGTCACAGAGGTGGTTTGGCTATCACTATCGCCACAGAGGTAAATGA 181
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  122 TCAAGAAAAAGGAATCAATGACATCATCATCGTGGTTATCTTAAAGAACAAATTCG 181
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  182 TTAATTTCTTGAGGACTTTCAGCATAAAGCTTGGCATCACAATACATGCTCCCAAGAGA 241
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  182 ATTAATTAAAGAAAAAATACGGTGTTCGCCCTCGTTTTCATGATAAATACGCTGACTACA 241
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  242 CTGA 245
    |||
DB  242 ATAA 245
    |||

RESULT 8
US-08-196-218-31
; Sequence 31, Application US/08196218
; Patent No. 5614619
; GENERAL INFORMATION:
; APPLICANT: Piepersberg, Wolfgang
; APPLICANT: Stockmann, Michael
; APPLICANT: Taleghani, Kampiz Mansouri
; APPLICANT: Distler, Jurgen
; APPLICANT: Grabley, Susanne
; APPLICANT: Sichel, Petra
; APPLICANT: Brau, Barbara
; TITLE OF INVENTION: Secondary-Metabolite Biosynthesis Genes
; TITLE OF INVENTION: From Actinomycetes, Method of Isolating Them, and Their
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSER: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/196,218
; FILING DATE: 25-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ogden, Stasia L.
; REGISTRATION NUMBER: 36,228
; REFERENCE/DOCKET NUMBER: 02481.1372-00000
; TELECOMMUNICATION INFORMATION:
; FILING DATE: 25-AUG-1994
```

```
/
/
/ TELEPHONE: 202-408-4000
/ TELEFAX: 202-408-4400
/ INFORMATION FOR SEQ ID NO: 31:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2634 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 3..401
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 416..1531
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 1561..2625
US-08-196-218-31

Query Match          4.9%; Score 52.8; DB 1; Length 2634;
Best Local Similarity 57.1%; Pred. No. 5.8e-07;
Matches 96; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY  1  ATGAAGCCCTCATCTTGTGCGGGGTTTCGGAACCCGCCCTTCGCGCTTTCAGCTCTGAGC 60
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  1561 ATGAAGGCTCTGGTCTCGCGCGGATCTGTTACCCGCTTGGGCTTTCAGTTATTTCG 1620
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  61  TTCCGGAACCCCTCGTGGATTTTGCACCAAGCCCATGATTCGCACAGATCGAAGCT 120
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  1621 ATGCCAAACAATCATCCCATGCCCAACACACCCGCTGCTGCTCATCTGCTGAACGCC 1680
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  121 TTGAAGAAGTTGGGTCACAGAGGTGGTTTGGCTATCAACTATCGC 168
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  1681 GTCCGGAGCTGGGCGTGACCGAGGTGCGGCTCATCTGCTGGCAACCGC 1728
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
US-08-681-953-31
; Sequence 31, Application US/08681953
; Patent No. 5710032
; GENERAL INFORMATION:
; APPLICANT: Piepersberg, Wolfgang
; APPLICANT: Stockmann, Michael
; APPLICANT: Taleghani, Kampiz Mansouri
; APPLICANT: Distler, Jurgen
; APPLICANT: Grabley, Susanne
; APPLICANT: Sichel, Petra
; APPLICANT: Brau, Barbara
; TITLE OF INVENTION: Secondary-Metabolite Biosynthesis Genes
; TITLE OF INVENTION: From Actinomycetes, Method of Isolating Them, and Their
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSER: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/681,953
; FILING DATE: 30-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/196,218
; FILING DATE: 25-AUG-1994
```

**QY** 235 CAAGAGACTGAGCCCTTAGGAACCGCTGGCCCTCTTGTCTCTAGCAAGGCACAAGCTTGCG 294  
||| || | ||| |||| | || |||| | ||| | | | | | | |  
**D6** 350 CAGGAATTTGCCCCCTAGSCACAGGGGGTGGTCTTTACCATTTCGACCCAGATCCTG 409

```
Query Match      3.8%; Score 41.6; DB 4; Length 6854;
Best Local Similarity 55.6%; Pred. No. 0.0049;
Matches 80: Conservative 0; Mismatches 64; Indels 0
```



```
; FILING DATE: 435
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/377,440
; FILING DATE: 24-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: GOLDBERG, JULES E.
; REGISTRATION NUMBER: 24,408
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-986-4090
; TELEFAX: 212-818-9479
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28804 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FRAGMENT TYPE: N-terminal
; US-08-592-874-1

Query Match          3.7%; Score 40.2; DB 2; Length 28804;
Best Local Similarity 45.0%; Pred. No. 0.033;
Matches 150; Conservative 0; Mismatches 183; Indels 0; Gaps 0;

QY 1 ATGAAGCCCTCATCTTGTGCGGGGTTTCGGAAACCCGCTTCGGCCCTTTGACTCTGAGC 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23238 ATGAAGGTATCATCTTGTGCGGGGCGAGCGGCGCTCTACCCGCAAGCTGTGC 23297

QY 61 TTCCGAAACCCCTCGTGGATTTCGAAACAGCCCATGATCTGCACCAAGATCGAAGCT 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23298 ATCTCGAACAGCTGCTTCCCGTCTATGACAAAGCCCATGATCTTACCCCTGTCCGTG 23357

QY 121 TTGAAAGAAAGTTGGGTCACAGAGTGGTTTGGCTATCAACTATGCCCGAGAGTAATG 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23358 CTGATGCTCAGGGTATCCGGGACATCCTGATCATCTCCACCCGCGGACCTGCCGATG 23417

QY 181 ATTAATTTCTGAAGGACTTTGAGGATAAGCTTGGGATACAAATTACATGCTCCCAAGAG 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23418 TTCCAGGCGCTGCTCGGCGACGGTTGCGCATTCGGCATCAACCTGAGCTATGCCGAACAG 23477

QY 241 ACTGAGCCCTTAGGAACCGCTGCCCTCTTGCTCTAGCAAGGACAAAGCTTGGCGGATGA 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23478 CCTTCGCCCAACGGCTTCGGGAAGCCTTCATCATCGGCGCGGATTTGCTCGCAACGAT 23537

QY 301 TCTGGCAGCATCTTTTGTCTCAACAGTGAT 333
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23538 CCCAGCGCTGATCTCGGCGGACACATCTAT 23570

RESULT 15
; Sequence 2, Application US/09096867
; Patent No. 6030817
; GENERAL INFORMATION:
; APPLICANT: Pollock, Thomas J
; APPLICANT: Mikolajczak, Marcia
; APPLICANT: Yamazaki, Motohide
; APPLICANT: Thorne, Linda
; APPLICANT: Armentrout, Richard W
; TITLE OF INVENTION: Production of Xanthan Gum by Sphingomonas Bacteria
; FILE REFERENCE: seq list for app'l filed from pro. appl
; CURRENT FILING DATE: 1998-06-11
; EARLIER APPLICATION NUMBER: 60/049,428
; EARLIER FILING DATE: 1997-06-12
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 28804
; TYPE: DNA
; ORGANISM: Sphingomonas sp. S88
; US-09-096-867-2

Query Match          3.7%; Score 40.2; DB 3; Length 28804;
Best Local Similarity 45.0%; Pred. No. 0.033;
Matches 150; Conservative 0; Mismatches 183; Indels 0; Gaps 0;

QY 1 ATGAAGCCCTCATCTTGTGCGGGGTTTCGGAAACCCGCTTCGGCCCTTTGACTCTGAGC 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23238 ATGAAGGTATCATCTTGTGCGGGGCGAGCGGCGCTCTACCCGCAAGCTGTGC 23297

QY 61 TTCCGAAACCCCTCGTGGATTTCGAAACAGCCCATGATCTGCACCAAGATCGAAGCT 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23298 ATCTCGAACAGCTGCTTCCCGTCTATGACAAAGCCCATGATCTTACCCCTGTCCGTG 23357

QY 121 TTGAAAGAAAGTTGGGTCACAGAGTGGTTTGGCTATCAACTATGCCCGAGAGTAATG 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23358 CTGATGCTCAGGGTATCCGGGACATCCTGATCATCTCCACCCGCGGACCTGCCGATG 23417

QY 181 ATTAATTTCTGAAGGACTTTGAGGATAAGCTTGGGATACAAATTACATGCTCCCAAGAG 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23418 TTCCAGGCGCTGCTCGGCGACGGTTTGGCATTCGGCATCAACCTGAGCTATGCCGAACAG 23477

QY 241 ACTGAGCCCTTAGGAACCGCTGGCCCTCTTGCTCTAGCAAGGACAAAGCTTGGCGGATGA 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23478 CCTTCGCCCAACGGCTTCGGGAAGCCTTCATCATCGGCGCGGATTTGCTCGCAACGAT 23537

QY 301 TCTGGCAGCATCTTTTGTCTCAACAGTGAT 333
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23538 CCCAGCGCTGATCTCGGCGGACACATCTAT 23570

RESULT 14
; Sequence 2, Application US/09096942
; Patent No. 6027925
; GENERAL INFORMATION:
; APPLICANT: Pollock, Thomas J
; APPLICANT: Mikolajczak, Marcia
; APPLICANT: Yamazaki, Motohide
; APPLICANT: Thorne, Linda
; APPLICANT: Armentrout, Richard W
; TITLE OF INVENTION: Production of Xanthan Gum by Sphingomonas Bacteria
; FILE REFERENCE: seq list for app'l filed from pro. appl
; CURRENT FILING DATE: 1998-06-12
; EARLIER APPLICATION NUMBER: 60/049,428
; EARLIER FILING DATE: 1997-06-12
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 28804
; TYPE: DNA
; ORGANISM: Sphingomonas sp. S88
```

```
US-09-096-942-2

Query Match          3.7%; Score 40.2; DB 3; Length 28804;
Best Local Similarity 45.0%; Pred. No. 0.033;
Matches 150; Conservative 0; Mismatches 183; Indels 0; Gaps 0;

QY 1 ATGAAGCCCTCATCTTGTGCGGGGTTTCGGAAACCCGCTTCGGCCCTTTGACTCTGAGC 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23238 ATGAAGGTATCATCTTGTGCGGGGCGAGCGGCGCTCTACCCGCAAGCTGTGC 23297

QY 61 TTCCGAAACCCCTCGTGGATTTCGAAACAGCCCATGATCTGCACCAAGATCGAAGCT 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23298 ATCTCGAACAGCTGCTTCCCGTCTATGACAAAGCCCATGATCTTACCCCTGTCCGTG 23357

QY 121 TTGAAAGAAAGTTGGGTCACAGAGTGGTTTGGCTATCAACTATGCCCGAGAGTAATG 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23358 CTGATGCTCAGGGTATCCGGGACATCCTGATCATCTCCACCCGCGGACCTGCCGATG 23417

QY 181 ATTAATTTCTGAAGGACTTTGAGGATAAGCTTGGGATACAAATTACATGCTCCCAAGAG 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23418 TTCCAGGCGCTGCTCGGCGACGGTTGCGCATTCGGCATCAACCTGAGCTATGCCGAACAG 23477

QY 241 ACTGAGCCCTTAGGAACCGCTGCCCTCTTGCTCTAGCAAGGACAAAGCTTGGCGGATGA 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23478 CCTTCGCCCAACGGCTTCGGGAAGCCTTCATCATCGGCGCGGATTTGCTCGCAACGAT 23537

QY 301 TCTGGCAGCATCTTTTGTCTCAACAGTGAT 333
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23538 CCCAGCGCTGATCTCGGCGGACACATCTAT 23570

RESULT 15
; Sequence 2, Application US/09096867
; Patent No. 6030817
; GENERAL INFORMATION:
; APPLICANT: Pollock, Thomas J
; APPLICANT: Mikolajczak, Marcia
; APPLICANT: Yamazaki, Motohide
; APPLICANT: Thorne, Linda
; APPLICANT: Armentrout, Richard W
; TITLE OF INVENTION: Production of Xanthan Gum by Sphingomonas Bacteria
; FILE REFERENCE: seq list for app'l filed from pro. appl
; CURRENT FILING DATE: 1998-06-11
; EARLIER APPLICATION NUMBER: 60/049,428
; EARLIER FILING DATE: 1997-06-12
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 28804
; TYPE: DNA
; ORGANISM: Sphingomonas sp. S88
; US-09-096-867-2

Query Match          3.7%; Score 40.2; DB 3; Length 28804;
Best Local Similarity 45.0%; Pred. No. 0.033;
Matches 150; Conservative 0; Mismatches 183; Indels 0; Gaps 0;

QY 1 ATGAAGCCCTCATCTTGTGCGGGGTTTCGGAAACCCGCTTCGGCCCTTTGACTCTGAGC 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23238 ATGAAGGTATCATCTTGTGCGGGGCGAGCGGCGCTCTACCCGCAAGCTGTGC 23297

QY 61 TTCCGAAACCCCTCGTGGATTTCGAAACAGCCCATGATCTGCACCAAGATCGAAGCT 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23298 ATCTCGAACAGCTGCTTCCCGTCTATGACAAAGCCCATGATCTTACCCCTGTCCGTG 23357

QY 121 TTGAAAGAAAGTTGGGTCACAGAGTGGTTTGGCTATCAACTATGCCCGAGAGTAATG 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23358 CTGATGCTCAGGGTATCCGGGACATCCTGATCATCTCCACCCGCGGACCTGCCGATG 23417

QY 181 ATTAATTTCTGAAGGACTTTGAGGATAAGCTTGGGATACAAATTACATGCTCCCAAGAG 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23418 TTCCAGGCGCTGCTCGGCGACGGTTTGGCATTCGGCATCAACCTGAGCTATGCCGAACAG 23477

QY 241 ACTGAGCCCTTAGGAACCGCTGCCCTCTTGCTCTAGCAAGGACAAAGCTTGGCGGATGA 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23478 CCTTCGCCCAACGGCTTCGGGAAGCCTTCATCATCGGCGCGGATTTGCTCGCAACGAT 23537

QY 301 TCTGGCAGCATCTTTTGTCTCAACAGTGAT 333
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23538 CCCAGCGCTGATCTCGGCGGACACATCTAT 23570
```

Db 23418 TTCCAGCGCGTCTCGGCGACGGTTTCGGGATTCGGGCATCAACCTGAGCTATGCCGAACAG 23477  
OY 241 ACTGAGCGCTTAGGACCGCGTGGCCCTCTGTCTCTAGCAAGGACAAAGCTTCGCCATGCA 300  
Db 23478 CCTCGCCCAACGGCCTTCGGCAAGCCTTCATCATCGGCGCGGATTCGTCGGCAGGAT 23537  
OY 301 TCTGGCCAGCCATTCTTTGTCTCTACACAGTGAT 333  
Db 23538 CCCAGCGCGCTGATCTCGGCGACACACATCTAT 23570

Search completed: November 27, 2002, 04:07:04  
Job time : 185 secs

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	187	17.2	212	10	US-09-294-093B-2164		Sequence 2164, App
2	151.4	13.9	364	10	US-09-878-574-2725		Sequence 2725, App
3	130.6	12.0	269	10	US-09-878-574-8673		Sequence 8673, App
4	79.2	7.3	272	10	US-09-878-574-10462		Sequence 10462, App
c	73.2	6.7	787	10	US-09-954-773A-20		Sequence 20, Appl
5	55.2	5.1	282	10	US-09-294-093B-3871		Sequence 3871, App
6	55.2	6.1	787	10	US-09-954-773A-20		Sequence 20, Appl
7	46.2	4.3	678	10	US-09-738-973-155		Sequence 155, App
c	38	3.5	167343	10	US-09-962-436-281		Sequence 281, App
8	38	3.5	167343	10	US-09-964-824A-273		Sequence 273, App
c	35.2	3.2	808	10	US-09-954-773A-19		Sequence 19, Appl
9	35.2	3.2	879	9	US-09-860-846-11		Sequence 11, Appl
10	35.2	3.2	879	9	US-09-861-289-11		Sequence 11, Appl
11	35.2	3.2	879	9	US-09-861-289-11		Sequence 11, Appl
12	35.2	3.2	879	9	US-09-861-289-11		Sequence 11, Appl
13	35.2	3.2	13613	9	US-09-860-846-3		Sequence 3, Appl
c	35.2	3.2	13613	10	US-09-861-289-3		Sequence 3, Appl
14	35.2	3.2	447	10	US-09-770-444-767		Sequence 767, App
15	34.8	3.2	396	10	US-09-878-574-757		Sequence 757, App
16	34.6	3.2	396	10	US-09-878-574-757		Sequence 757, App
c	34.4	3.2	1130	10	US-09-939-980-248		Sequence 248, App
17	34.4	3.2	1130	10	US-09-939-980-248		Sequence 248, App
18	33.8	3.1	410	10	US-09-878-574-380		Sequence 380, App
19	33.6	3.1	13808	10	US-09-070-927A-271		Sequence 271, App

DD 122 ACATTACNGGC-TGGCTCTTTATCTAGACTCGATTAGGAAGAAATCAGCTGCCAAGCTAG 180  
QY 737 CTACTGACACATGTTGTGGCAATGTGCTG 768  
DB 181 CTACTGACACATGTTGTGGCAATGTGCTG 212

RESULT 2

US-09-878-574-2725  
; Sequence 2725, Application US/09878574  
; Patent No. US20020110548A1  
; GENERAL INFORMATION:  
; APPLICANT: Byrum, Joseph R.  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Thompson, Michael D.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(15401)B  
; CURRENT APPLICATION NUMBER: US/09/878,574  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 09/333,535  
; PRIOR FILING DATE: 1999-06-14  
; NUMBER OF SEQ ID NOS: 15775  
; SEQ ID NO 2725  
; LENGTH: 364  
; TYPE: DNA  
; ORGANISM: Glycine max  
; OTHER INFORMATION: Clone ID: LIB3028-021-Q1-B1-C7  
US-09-878-574-2725

Query Match 13.9%; Score 151.4; DB 10; Length 364;  
Best Local Similarity 74.3%; Pred. No. 1.5e-39;  
Matches 191; Conservative 0; Mismatches 66; Indels 0; Gaps 0;  
QY 1 ATGAAGCCCTCATCTTCTCGGGGGTTTCGGAAACCCGCTTCGGCCTTTGACTCTGAGC 60  
DB 108 ATGAAGCCATGATTTCTGTTGGGGATTGGAACAAGCTGAGGCCACTGACACTCACT 167  
QY 61 TTCCCGAAACCCCTCGTGCATTTTGCACAAAGCCCATGATCTGCACCAAGATCGAAGCT 120  
DB 168 TTCCCTAAGCCCTCTGTTGATTTGCTACAGCCCTATGATTTTGGATCAGATAGAGCC 227  
QY 121 TTGAAGAAGATTTGGGTACAGAGGTGGTTTGGCTATCAACTATATGCCCGAGAGGTAATG 180  
DB 228 CTTAAGGCCATTGGAGTCTGAGTGTGCTAGCCATCAATTTACCAACAGAGGTTATG 287  
QY 181 ATTAATTTCTGAAGACATTTGAGGATAGCTTGGCATCACAAATTCATCTCCCAAGAG 240  
DB 288 TTGAATTTCTGAAGGATTTTGAATCAAGCTCGGCATCAAGATCACATGTTCTCAGGA 347  
QY 241 ACTGAGCCCTTAGGAC 257  
DB 348 ACTGAACCACTGGGAAC 364

RESULT 3

US-09-878-574-9673  
; Sequence 8673, Application US/09878574  
; Patent No. US20020110548A1  
; GENERAL INFORMATION:  
; APPLICANT: Byrum, Joseph R.  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Thompson, Michael D.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(15401)B  
; CURRENT APPLICATION NUMBER: US/09/878,574  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 09/333,535  
; PRIOR FILING DATE: 1999-06-14  
; NUMBER OF SEQ ID NOS: 15775  
; SEQ ID NO 8673  
; LENGTH: 269

TVDE: DNA  
; ORGANISM: Glycine max  
; OTHER INFORMATION: Clone ID: 701101510H1  
US-09-878-574-8673  
Query Match 12.0%; Score 130.6; DB 10; Length 269;  
Best Local Similarity 70.3%; Pred. No. 7.9e-33;  
Matches 189; Conservative 0; Mismatches 79; Indels 1; Gaps 1;  
QY 230 GCTCCCAAGACACTGACCCCTTAGGAACCGCTGGCCCTCTTGTCTCTAGCAAGGACAAGC 289  
DB 1 GTTCTCAGGAACCTGAAACCATTTGGGAACAGCAGAGTCTCTGGCTCTTCCCAAGGATAAGC 60  
QY 290 TTGGGATGATCTGGCC-ACCATTCTTTCTCTCAACACTGATGTCATTAAGCGAATAC 348  
DB 61 TCATATACATCTCTGAGAACCCCTTTTGTTCACACAGTGTATATCAGTGAAT 120  
QY 349 CCATTTCGTGAACCTCATCAAAATTCACAAGTGTCTATGGTGTGAGGCAACAATATATGTC 408  
DB 121 CCACTCAAGAAATGATTAATCCATACAAACCAGGAGAGGCTTCCATATATGTA 180  
QY 409 ACTAAGCTGGATGAACCATCAAAATACGGTGTGTGTGTTATGAGGAGGCAACTGGCAGG 468  
DB 181 ACCAAGTTGACGAGCCATCAAGTACGGCTGCTGTGATGGAAGACACACAGGCCAG 240  
QY 469 GTGGAAAGTTTGTGTGAGAGCCAAAAT 497  
DB 241 GTTGATAAATTTGTTGAACCAACCGAAAT 269

RESULT 4  
US-09-878-574-10462  
; Sequence 10462, Application US/09878574  
; Patent No. US20020110548A1  
; GENERAL INFORMATION:  
; APPLICANT: Byrum, Joseph R.  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Thompson, Michael D.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(15401)B  
; CURRENT APPLICATION NUMBER: US/09/878,574  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 09/333,535  
; PRIOR FILING DATE: 1999-06-14  
; NUMBER OF SEQ ID NOS: 15775  
; SEQ ID NO 10462  
; LENGTH: 272  
; TYPE: DNA  
; ORGANISM: Glycine max  
; OTHER INFORMATION: Clone ID: 700966702H1  
US-09-878-574-10462

Query Match 7.3%; Score 79.2; DB 10; Length 272;  
Best Local Similarity 72.9%; Pred. No. 5.5e-16;  
Matches 102; Conservative 0; Mismatches 38; Indels 0; Gaps 0;  
QY 1 ATGAAGCCCTCATCTTCTGTCGGGGTTTCGGAACCGCCCTTCGGCCTTTGACTCTGAGC 60  
DB 132 ATGAAGCATTTGATCTGTGTTGGGGTTTGGAAACAAGGTTGAGGCCCACTCAGACTCAGT 191  
QY 61 TTCCCGAAACCCCTCGTGCATTTTGCACAAAGCCCATGATTCGCAACAGATGGAAGCT 120  
DB 192 TTCTCTAGGCTCTGTGTGATTTTGCYACAAACCAACCTATGATTCGTCATCAGATAGAGCC 251  
QY 121 TTGAAGAAGTGTGGGTCAC 140  
DB 252 CTTAAGGCCCATTTGAGTTAC 271

RESULT 5  
US-09-954-773A-20/c  
; Sequence 20, Application US/09954773A



```
; Patent No. US20020129402A1
; GENERAL INFORMATION:
; APPLICANT: Southern Illinois University
; APPLICANT: Lightfoot, David A
; APPLICANT: Gibson, Paul T
; APPLICANT: Meksem, Khalid
; TITLE OF INVENTION: SOYBEAN SUDDEN DEATH SYNDROME RESISTANT SOYBEANS, SOYBEAN CYST NE
; FILE REFERENCE: 1268/2/2
; CURRENT APPLICATION NUMBER: US/09/954,773A
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: 09/007,119
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 787
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(787)
; OTHER INFORMATION: n is a, c, g, or t/u
US-09-954-773A-20

Query Match 6.7%; Score 73.2; DB 10; Length 787;
Best Local Similarity 56.0%; Pred. No. 1.1e-13;
Matches 227; Conservative 0; Mismatches 151; Indels 27; Gaps 5;

QY 524 GGATTACTTACTGAACCATCTGCTCTGACGGCATTGAGCTGAGGCCCAACATCAATG 583
Db 563 GGGTAACTGGATGAACCCCTTGATGGANAGATNGAGNGACACCNATTTCTTGA 504
QY 584 AGAAGAGCTCTCCCTCAAAATGACGTCATCAACAGCTCTATGCAATGGCTTCC-- 641
Db 503 AGACAGAGGGTTTCAAAATATGTTTCNGATAGAAAGCTGNATGCAATGGCTTTCCAC 444
QY 642 ----AGTGTGTTGGATGATGTTGTCAGCTAGGAGCTACATCTACTGCTGCTTT 697
Db 443 AGGAAGATTNAGANGAAATNGGACAA---AGGACTATATTTNNGANTGAGACTCN 387
QY 698 ATCTAGACTGATTAGGAAGA---AATCAGCTGCCAAGCTAGCTAGGAGCACATGTTG 754
Db 386 ACNTGAACCTGGGAGGAATAATNTCTTAGTGAACNGGCCAATGACNTCANACATANTG 327
QY 755 TTGGCAATGCTGTTGTCATGAGAGCGCCCAAGATTTGGAGAAGTTGCTGATTTGCTG 814
Db 326 TGGGAAATGTTGTTGTCATGAGACTGCCAA-ATNGGTGATGATGTCATAAATTTGGTCC-- 270
QY 815 ATGTGCCATTGGACCTGGTGTGTTGTGGAGGACGCGTGGAGCTTTCCCGCTGCACTG 874
Db 269 -----ATCNGGATGATCATTTGAGAAGAGTTGGGCTCAAAATCTNACACTG 222
QY 875 TCATCGCGCGCTGCTATCAAGAAGCATGCTTGCATCTCAACA 919
Db 221 TTATGTGAAGCACTCGTGTATAAAGCAATTTGTTATATCNGACA 177

RESULT 6
US-09-294-093B-3871
; Sequence 3871, Application US/09294093B
; Patent No. US20010051335A1
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath, V.
; APPLICANT: Ito, Laura, Y.
; APPLICANT: Sherman, Bradley, K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
; FILE REFERENCE: PL-0009 US
; CURRENT APPLICATION NUMBER: US/09/294,093B
; PRIOR FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 60/082,567
; PRIOR FILING DATE: April 21, 1998
; NUMBER OF SEQ ID NOS: 6207
```

```
; SOFTWARE: PERL Program
; SEQ ID NO 3871
; LENGTH: 282
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20010051335A1 700352730H1
; NAME/KEY: unsure
; LOCATION: 50, 185, 218
; OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-3871

Query Match 5.1%; Score 55.2; DB 10; Length 282;
Best Local Similarity 89.9%; Pred. No. 4.1e-08;
Matches 71; Conservative 0; Mismatches 4; Indels 4; Gaps 1;

QY 1 ATGAAGGCCCTCATCTTGTCTGGGGTTTCGGAACCCGCTTCGGCCTTTGACTCTGAGC 60
Db 208 ATGAAGGCCCNCAATCTTGTCTGGGGTTTCGGAACCCGCTTCGGCCTTAGATT----GG 263
QY 61 TTCCCGGAACCCCTCGTGG 79
Db 264 TTCCCGGAACCCCTCGTGG 282

RESULT 7
US-09-738-973-155
; Sequence 155, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Fling, Steven P.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Algate, Paul A.
; APPLICANT: Secrist, Heather
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Benson, Darin R.
; APPLICANT: Eliot, Mark
; APPLICANT: Mannion, Jane
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 21021.475C9
; CURRENT APPLICATION NUMBER: US/09/738,973
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 155
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-738-973-155

Query Match 4.3%; Score 46.2; DB 10; Length 678;
Best Local Similarity 49.5%; Pred. No. 6.7e-05;
Matches 151; Conservative 0; Mismatches 148; Indels 6; Gaps 1;

QY 235 CAAGAGACTGAGCCCTTAGGACGCTGGCCCTCTTGTCTAGCAAGGACAAGCTTGGC 294
Db 350 CAGGAATTTGCCCCCTTAGGCACAGGGGTGCTCTTACCATTTCGAGACAGATCTCTG 409
QY 295 GATGGATCTGGCCAGCCATCTTTGTCTCAACAGTGTATGTCATAAGCGAATACCCATTT 354
Db 410 GCTGGGAGCCCGGAGGATCTTCTGCTCAATGCTGATGCTGCTCGACTTCCCTTG 469
QY 355 GCTGAATCTATCAAAATTTCACAAGTGTATGTTGGTGGAGCAACAATTTATGTTCACTAAG 414
Db 470 AGTGTATGTTGGAAGCCCGACCGCGCTCACCCCTTTCTTACTCTTGGCACTAGC 529
QY 415 GTGGATG-----AACCATCAAAATACGGTGTGTGGTGTATGGAGGAGGCAACTGGCAGG 468
```

Db 530 GCTAACAGGACGCAATCCCTCACTACCTGGCTGCATGCTTGAGATCCACAGACACACGAG 589  
QY 469 GTGGAAGGTTTGTGAGAGCCCAAAATATTTGTGGTAAACAGATCAATGCTGGGATT 528  
Db 590 GTATTGACATGCTGGAGAAACCCACACATTTATCAGTCACATCATCACTGGGGACC 649  
QY 529 TACTT 533  
Db 650 TACCT 654

## RESULT 8

US-09-962-436-281/c  
; Sequence 281, Application US/09962436  
; Patent No. US20020081301a1  
; GENERAL INFORMATION:  
; APPLICANT: Soppet, Daniel  
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu  
; FILE REFERENCE: 689290-75  
; CURRENT APPLICATION NUMBER: US/09/962,436  
; CURRENT FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,082  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US/60/234,924  
; PRIOR FILING DATE: 2000-09-25  
; NUMBER OF SEQ ID NOS: 568  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 281  
; LENGTH: 167343  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-962-436-281

Query Match 3.5%; Score 38; DB 10; Length 167343;  
Best Local Similarity 52.5%; Pred. No. 1.4;  
Matches 83; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 929 GCTGGCACCACACTGTGGTCAATGGCCAGGATAGAGATATGACATATCCTGGGGAGG 988  
Db 120438 GCCAGGATTTCAGTGTAAAGGATGGCCAGGAGAGTATGGGATTGGGAACGAG 120379  
QY 989 ATGTCATGTGTGATAGGTGTACAGCAATGGCGGTGTGTCTCCACATAAAGAGA 1048  
Db 120378 AATGGGATGTAGCTGTCAAACTAGACGATCACTGCTAGACATCTTACAAATAAACCCT 120319  
QY 1049 TCAAGTCAGCATCTCTGAAGCCTGAGATCGTCATGTGA 1086  
Db 120318 TTAAGCTTGGAAACTGTTGAATGTGAAGGTAAGAGA 120281

## RESULT 9

US-09-964-824A-273/c  
; Sequence 273, Application US/09964824A  
; Patent No. US20020102531a1  
; GENERAL INFORMATION:  
; APPLICANT: Horrigan, Stephen  
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu  
; FILE REFERENCE: 689290-73  
; CURRENT APPLICATION NUMBER: US/09/964,824A  
; CURRENT FILING DATE: 2001-09-27  
; PRIOR APPLICATION NUMBER: US/60/236,033  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US/60/236,032  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US/60/236,028  
; PRIOR FILING DATE: 2000-09-28  
; NUMBER OF SEQ ID NOS: 583  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 273  
; LENGTH: 167343

; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-964-824A-273  
Query Match 3.5%; Score 38; DB 10; Length 167343;  
Best Local Similarity 52.5%; Pred. No. 1.4;  
Matches 83; Conservative 0; Mismatches 75; Indels 0; Gaps 0;  
QY 929 GCTGGCACCACACTGTGGTCAATGGCCAGGATAGAGATATGACATATCCTGGGGAGG 988  
Db 120438 GCCAGGATTTCAGTGTAAAGGATGGCCAGGAGAGTATGGGGATTGGGAACGAG 120379  
QY 989 ATGTCATGTGTGATAGGTGTACAGCAATGGCGGTGTGTCTCCACATAAAGAGA 1048  
Db 120378 AATGGGATGTAGCTGTCAAACTAGAGATCACTGCTAGACATCTTACAAATAAACCCT 120319  
QY 1049 TCAAGTCAGCATCTCTGAAGCCTGAGATCGTCATGTGA 1086  
Db 120318 TTAAGCTTGGAAACTGTTGAATGTGAAGGTAAGAGA 120281

## RESULT 10

US-09-954-773A-19/c  
; Sequence 19, Application US/09954773A  
; Patent No. US20020129402a1  
; GENERAL INFORMATION:  
; APPLICANT: Southern Illinois University  
; APPLICANT: Lightfoot, David A  
; APPLICANT: Gidson, Paul T  
; APPLICANT: Meksem, Khalid  
; TITLE OF INVENTION: SOYBEAN SUDDEN DEATH SYNDROME RESISTANT SOYBEANS, SOYBEAN CYST  
; FILE REFERENCE: 1268/2/2  
; CURRENT APPLICATION NUMBER: US/09/954,773A  
; CURRENT FILING DATE: 2002-01-22  
; PRIOR APPLICATION NUMBER: 09/007,119  
; PRIOR FILING DATE: 1998-01-14  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 19  
; LENGTH: 808  
; TYPE: DNA  
; ORGANISM: Glycine max  
; NAME/KEY: misc feature  
; LOCATION: (1)-(808)  
; OTHER INFORMATION: n is a, c, g, or t/u  
US-09-954-773A-19

Query Match 3.2%; Score 35.2; DB 10; Length 808;  
Best Local Similarity 46.7%; Pred. No. 0.3;  
Matches 196; Conservative 0; Mismatches 201; Indels 23; Gaps 4;

QY 518 ATGCTGGGATTTACTTACTGAACCCATCTGTCTTGGACCGCATTT-GAGCTGAGGCCACA 576  
Db 707 ATNCGGGTTAAAGTGGAGGAACCCCTTNGGAATGGANAGAAATGGAGNTGAACCAATT 648  
QY 577 TCAATTCAGAAAGAGGCTTCCCTCAAAATTCAGCTGATCAACAGCTCTATCAATGTC 636  
Db 647 TCTTNGAAACAGAGGGTTTCCNAAAAATNGNTTTGGGAAGAANGCTGGAAGCAAAANGT 588  
QY 637 CTTCAGGTTTTTGGG-----TGGATGTGTGTCAGCTAGGAGCTACATCTACTGCGCTGCG 692  
Db 587 CTTTTCNAACAGNANGAATTTGNAGATGGAAATGGCAACGGGATATATTTTNNNGGATGAG 528  
QY 693 TCTTTATCTAGACTCGATTAGGAAGAAATCAGCTGCCAAGCTAGCTACTGAGAGCACA--- 749  
Db 527 ANTCAACCTGAANNCGGCGGAGCAATATTTNTAAGGAACNGCGCAANGACCTTCAAAAC 468  
QY 750 TCTTGTGGCAATGCTGCTGCTGCATGAGAGCGCCCAAGATTGGAGAGGTTGCTCTGATTGG 809  
Db 467 ATACTGNGGGAATGTGTGCTGCATGAGACTNCCAAATNGTGGAGNNATGTCAAAANGT 408

QY 810 TCCTGATGTCGCCATTGACCTGGTGTGTTGTGAGAGACGGCGTGAGGCTTTCCCGCTG 869  
Db 407 TCAT-----CCTGNATGATCATTTGAGAAGGAGNCTGGCTCAAATCTAA 363  
QY 870 CACTGTGATCGCGCGGTGCTATCAAGAAGCATGCTTGCATCTCAAAACAGCATTTATCGG 929  
Db 362 CANTGTNAGGGGAAGCACTCGTGTATAAAGCAATTTNGTTATATCNGACAATGNGATTNG 303

RESULT 11

US-09-860-846-11  
; Sequence 11, Application US/09860846  
; Patent No. US20020164742A1  
; GENERAL INFORMATION:  
; APPLICANT: Sherman, D.H.  
; APPLICANT: Liu, H.  
; APPLICANT: Xue, Y.  
; APPLICANT: Zhao, L.  
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin  
; FILE REFERENCE: 600.438US1  
; CURRENT APPLICATION NUMBER: US/09/860.846  
; CURRENT FILING DATE: 2001-05-18  
; PRIOR APPLICATION NUMBER: 09/105,537  
; PRIOR FILING DATE: 1998-06-26  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 11  
; LENGTH: 879  
; TYPE: DNA  
; ORGANISM: Streptomyces venezuelae  
US-09-860-846-11

Query Match 3.2%; Score 35.2; DB 9; Length 879;  
Best Local Similarity 44.2%; Pred. No. 0.32;  
Matches 145; Conservative 0; Mismatches 183; Indels 0; Gaps 0;

QY 1 ATGAAGCCCTCATCTCTGTGCGGGGTTTCGGAACCCGCTTCGGCCCTTTGACTCTGAGC 60  
Db 1 ATGAAGGAATAGTCTCTGCGCGGGGAGCGAACTCGGCTGCATCGCGGACCTCGGTC 60  
QY 61 TTCCGAAACCCCTCGTGGATTTCGAAACAGCCCATGATTCGACCCAGATCGAAAGCT 120  
Db 61 ATTTGGAAGCAGATCTCTCCGGTCTACAACAAACCGATGATCTACTATCCGCTGTCGGTT 120  
QY 121 TTGAAGAAGTTGGGTCACAGAGTGTTTGGCTATCACTATCGCCCGAGAGTAATG 180  
Db 121 CTCATGCTCGCGGTATTCGGAGATTCAAATCATCTCGACCCCGCAGCATCGAACTC 180  
QY 181 ATTAATTTCTGAAGGACTTTGAGGATAAGCTTGGCATCACAATTACATGCTCCCAAG 240  
Db 181 TTCCAGTCGCTTCTCGGAACGGCAGGCACCTGGGAATAGAACTGACTATCGGTCAG 240  
QY 241 ACTGAGCCCTTAGAACCGCTGGCCCTTCGGAACCCGCTTCGGCCCTTTGACTCTGAGC 300  
Db 241 AAAGAGCCCGAGGAATCGCGGCGGCACTTCTCGTGGAGCGGAGCATCGCCCGCAGCATCG 300  
QY 301 TCTGCGCAGCATCTTGTCTCTCAACA 328  
Db 301 ACCTGCGCCGTGATCTCTGGGCGACAACA 328

RESULT 12

US-09-861-289-11  
; Sequence 11, Application US/09861289  
; Patent No. US20020110897A1  
; GENERAL INFORMATION:  
; APPLICANT: Sherman, D.H.  
; APPLICANT: Liu, H.  
; APPLICANT: Xue, Y.  
; APPLICANT: Zhao, L.  
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin  
; FILE REFERENCE: 600.438US1  
; CURRENT APPLICATION NUMBER: US/09/861.289

; CURRENT FILING DATE: 2001-05-18  
; PRIOR APPLICATION NUMBER: 09/105,537  
; PRIOR FILING DATE: 1998-06-26  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 11  
; LENGTH: 879  
; TYPE: DNA  
; ORGANISM: Streptomyces venezuelae  
US-09-861-289-11

Query Match 3.2%; Score 35.2; DB 10; Length 879;  
Best Local Similarity 44.2%; Pred. No. 0.32;  
Matches 145; Conservative 0; Mismatches 183; Indels 0; Gaps 0;

QY 1 ATGAAGCCCTCATCTCTGTGCGGGGTTTCGGAACCCGCTTCGGCCCTTTGACTCTGAGC 60  
Db 1 ATGAAGGAATAGTCTCTGCGCGGGGAGCGAACTCGGCTGCATCGCGGACCTCGGTC 60  
QY 61 TTCCGAAACCCCTCGTGGATTTCGAAACAGCCCATGATTCGACCCAGATCGAAAGCT 120  
Db 61 ATTTGGAAGCAGATCTCTCCGGTCTACAACAAACCGATGATCTACTATCCGCTGTCGGTT 120  
QY 121 TTGAAGAAGTTGGGTCACAGAGTGTTTGGCTATCACTATCGCCCGAGAGTAATG 180  
Db 121 CTCATGCTCGCGGTATTCGGAGATTCAAATCATCTCGACCCCGCAGCATCGAACTC 180  
QY 181 ATTAATTTCTGAAGGACTTTGAGGATAAGCTTGGCATCACAATTACATGCTCCCAAG 240  
Db 181 TTCCAGTCGCTTCTCGGAACGGCAGGCACCTGGGAATAGAACTGACTATCGGTCAG 240  
QY 241 ACTGAGCCCTTAGAACCGCTGGCCCTTCGCTCTAGCAAGGACAAGCTTGGGGATGGA 300  
Db 241 AAAGAGCCCGAGGAATCGCGGCGGCACTTCTCGTGGAGCGGAGCATCGCCCGCAGCATCG 300  
QY 301 TCTGCGCAGCATCTTGTCTCTCAACA 328  
Db 301 ACCTGCGCCGTGATCTCTGGGCGACAACA 328

RESULT 13

US-09-860-846-3/c  
; Sequence 3, Application US/09860846  
; Patent No. US20020164742A1  
; GENERAL INFORMATION:  
; APPLICANT: Sherman, D.H.  
; APPLICANT: Xue, Y.  
; APPLICANT: Zhao, L.  
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin  
; FILE REFERENCE: 600.438US1  
; CURRENT APPLICATION NUMBER: US/09/860.846  
; CURRENT FILING DATE: 2001-05-18  
; PRIOR APPLICATION NUMBER: 09/105,537  
; PRIOR FILING DATE: 1998-06-26  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3  
; LENGTH: 13613  
; TYPE: DNA  
; ORGANISM: Streptomyces venezuelae  
US-09-860-846-3

Query Match 3.2%; Score 35.2; DB 9; Length 13613;  
Best Local Similarity 44.2%; Pred. No. 2.1;  
Matches 145; Conservative 0; Mismatches 183; Indels 0; Gaps 0;

QY 1 ATGAAGCCCTCATCTCTGTGCGGGGTTTCGGAACCCGCTTCGGCCCTTTGACTCTGAGC 60  
Db 9989 ATGAAGGAATAGTCTCTGCGCGGGGAGCGAACTCGGCTGCATCGCGGACCTCGGTC 9930  
QY 61 TTCCGAAACCCCTCGTGGATTTCGAAACAGCCCATGATTCGACCCAGATCGAAAGCT 120

Db 9929 ATTTCGAGCAGATCTTCGGGTCTACACAAACCCGATGATCTACTATCCGCTGTTCGGTT 9870  
QY 121 TTGAAAGAAGTTGGGTACAGAGTGGTTTGGCTATCAACTATCGCCACAGAGTAATG 180  
Db 9869 CTCATGCTCGGCGGTATTCCGAGATTCAAATCATCTCGACCCCGACATCGAACTC 9810  
QY 181 ATTAATTTCTTGAAGACATTTAGGATATAGCTTGGCATCACAATTACATGCTCCCAAGAG 240  
Db 9809 TTCCAGTCCGCTTCTCGGAACGCCACGCCCTGGGAATAGAACTCGACTATCGGTCAG 9750  
QY 241 ACTGAGCCCTTAGGAACCGCTGCCCTCTTCTTCTAGCAAGGACAAAGCTTGGCGGATGCA 300  
Db 9749 AAGAGCCCGCAGGATCGGAGCGACATCTCTGTCGGAGCCGAGCACAATCGCGGACGAC 9690  
QY 301 TCTGGCCAGCCATTCCTTCTCTCTCAACA 328  
Db 9689 ACCTGGCCCTGATCTCTGGCGACACA 9662

## RESULT 14

US-09-861-289-3/c  
; Sequence 3, Application US/09861289  
; Patent No. US20020110897A1  
; GENERAL INFORMATION:  
; APPLICANT: Sherman, D.H.  
; APPLICANT: Liu, H.  
; APPLICANT: Xue, Y.  
; APPLICANT: Zhao, L.  
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin  
; FILE REFERENCE: 600.438US1  
; CURRENT APPLICATION NUMBER: US/09/861,289  
; CURRENT FILING DATE: 2001-05-18  
; PRIOR APPLICATION NUMBER: 09/105,537  
; PRIOR FILING DATE: 1998-06-26  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: FASTSEQ for Windows Version 3.0  
; SEQ ID NO 3  
; LENGTH: 13613  
; TYPE: DNA  
; ORGANISM: Streptomyces venezuelae  
US-09-861-289-3

Query Match 3.2%; Score 35.2; DB 10; Length 13613;  
Best Local Similarity 44.2%; Pred. No. 2.1;  
Matches 145; Conservative 0; Mismatches 183; Indels 0; Gaps 0;

QY 1 ATGAGAGCCCTCATCTTCTCGGGGTTCGGAACCCGCTTCGGCCTTTGACTCTTGAGC 60  
Db 9989 ATCAAGCGAATATCTCTGCGCGCGGAGCGGAACTCGCTGCATCCGGCGACCTCGGT 9930  
QY 61 TTCCCGARACCCCTCTGTGATTTTGCAAACAGCCCATCATCTTGCACCAAGATCGAAGCT 120  
Db 9929 ATTTCGAGCAGATCTTCGGGTCTACAAACAAACCGATGATCTACTATCCGCTGTTCGGTT 9870  
QY 121 TTGAAAGAAGTTGGGTACAGAGTGGTTTGGCTATCAACTATCGCCACAGAGTAATG 180  
Db 9869 CTCATGCTCGGCGGTATTCCGAGATTCAAATCATCTCCACCCCGACATCGAACTC 9810  
QY 181 ATTAATTTCTTGAAGACATTTAGGATAAGCTTGGCATCACAATTACATGCTCCCAAGAG 240  
Db 9809 TTCCAGTCCGCTTCTCGGAACGCCACGCCCTGGGAATAGAACTCGACTATCGGTCAG 9750  
QY 241 ACTGAGCCCTTAGGACCGCTGGCCTCTTGTCTTAGCAAGGAGACAAGCTTGGCGGATGCA 300  
Db 9749 AAGAGCCCGCAGGATCGGAGCGACATCTCTGTCGGAGCCGAGCACAATCGCGGACGAC 9690  
QY 301 TCTGGCCAGCCATTCCTTCTCTCTCAACA 328  
Db 9689 ACCTGGCCCTGATCTCTGGCGACACA 9662

## RESULT 15

US-09-770-444-767

\* - \*

; Sequence 767, Application US/09770444  
; Patent No. US20020023280A1  
; GENERAL INFORMATION:  
; APPLICANT: Gotlach, Jorn  
; APPLICANT: An, Yong-Qiang  
; APPLICANT: Hamilton, Carol M.  
; APPLICANT: Price, Jennifer L.  
; APPLICANT: Raines, Tracy M.  
; APPLICANT: Yu, Yang  
; APPLICANT: Rameaka, Joshua G.  
; APPLICANT: Page, Amy  
; APPLICANT: Matthew, Abraham V.  
; APPLICANT: Ledford, Brooke L.  
; APPLICANT: Woessner, Jeffrey P.  
; APPLICANT: Haas, William David  
; APPLICANT: Garcia, Carlos A.  
; APPLICANT: Krieker, Maja  
; APPLICANT: Slader, Ted  
; APPLICANT: Davis, Keith R.  
; APPLICANT: Allen, Keith  
; APPLICANT: Hoffman, Neil  
; APPLICANT: Hurdan, Patrick  
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis  
; TITLE OF INVENTION: thaliana  
; FILE REFERENCE: 2027 (PARA-016PRV)  
; CURRENT APPLICATION NUMBER: US/09/770,444  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 60/178,502  
; PRIOR FILING DATE: 2000-01-27  
; NUMBER OF SEQ ID NOS: 999  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 767  
; LENGTH: 447  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-770-444-767

Query Match 3.2%; Score 34.8; DB 10; Length 447;  
Best Local Similarity 58.8%; Pred. No. 0.27;  
Matches 60; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 31 CGAACCCGCTTCGGCCTTTGACTCTGAGCTTCCGAAACCCCTCGTGGATTTGCAAC 90  
Db 135 GGCACTCGATTCGGCCATTTCGCTGAATATCCAAAGCCTCTGTTCCCTATTGCGGA 194  
QY 91 AAGCCCATGATTCGACCAAGATCGAAGCTTTGAAAGAGTT 132  
Db 195 CAACCAATGTCATCATCATTCAGCTTGTAAGAAAT 236

Search completed: November 27, 2002, 05:08:44  
Job time : 353 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2002, 02:38:36 ; Search time 1981 Seconds  
(without alignments)  
8878.495 Million cell updates/sec

Title: US-09-374-967-1

Perfect score: 1086

Sequence: 1 atgaaggccctattcttgt.....agcctgagatgctatgtga 1086

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estopl:\*

7: em\_estro:\*

8: em\_esthc:\*

9: gb\_estl:\*

10: gb\_est2:\*

11: gb\_est3:\*

12: gb\_est4:\*

13: gb\_est5:\*

14: gb\_est6:\*

15: em\_estfun:\*

16: em\_estom:\*

17: gb\_gss:\*

18: em\_gss\_hum:\*

19: em\_gss\_inv:\*

20: em\_gss\_pln:\*

21: em\_gss\_vrt:\*

22: em\_gss\_fun:\*

23: em\_gss\_mam:\*

24: em\_gss\_mus:\*

25: em\_gss\_other:\*

26: em\_gss\_pro:\*

27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1068.4	98.4	1542	11 AY108204	AY108204 Zea mays
2	753.8	69.4	980	12 BG321249	BG321249 Zm04_04b0
3	538.2	49.6	694	12 BG049329	BG049329 OVL18_E1
4	490.4	45.2	836	10 BE455732	BE455732 HVSMEg001
5	486	44.8	524	13 BM324237	BM324237 PIC1_26_A
6	483.4	44.5	597	10 BE405374	BE405374 WHE1213_F

c	7	478.6	44.1	686	13	BJ320079	BJ320079
c	8	476.4	43.9	802	12	BG582583	BG582583
c	9	460.2	42.4	697	13	BJ318153	BJ318153
c	10	459.4	42.3	654	13	BI075102	BI075102
c	11	447.4	41.2	651	10	AV935887	AV935887
c	12	442.8	40.8	794	12	BE844694	BE844694
c	13	438	40.3	619	14	BQ806351	BQ806351
c	14	437.4	40.3	673	10	AV928662	AV928662
c	15	435	40.1	808	12	BG414225	BG414225
c	16	434.4	40.0	774	13	BM413164	BM413164
c	17	433	39.9	789	13	BI421941	BI421941
c	18	431.6	39.7	474	13	BM499735	BM499735
c	19	427	39.3	612	10	AV925892	AV925892
c	20	421.4	38.8	794	12	BG351126	BG351126
c	21	420	38.7	684	14	BQ970876	BQ970876
c	22	417.4	38.4	870	12	BF254904	BF254904
c	23	415.2	38.2	609	10	AV933434	AV933434
c	24	414.4	38.2	698	17	BH875336	BH875336
c	25	407.2	37.5	642	13	BQ265250	BQ265250
c	26	396.4	36.5	687	14	BQ506346	BQ506346
c	27	396	36.5	665	13	BJ314593	BJ314593
c	28	391	36.0	623	14	BM779448	BM779448
c	29	390.6	36.0	650	10	AM692513	AM692513
c	30	387.6	35.7	618	14	BM779156	BM779156
c	31	387.4	35.7	706	11	AY105397	AY105397
c	32	385.6	35.5	658	14	BQ789419	BQ789419
c	33	384	35.4	674	12	BG587608	BG587608
c	34	381.8	35.2	651	10	AM287855	AM287855
c	35	381	35.1	669	14	BQ487033	BQ487033
c	36	381	35.1	729	13	BI933297	BI933297
c	37	378.6	34.9	694	14	BQ743338	BQ743338
c	38	378.2	34.8	774	13	BI932250	BI932250
c	39	377.6	34.8	639	13	BJ259364	BJ259364
c	40	372.6	34.3	679	12	BG350520	BG350520
c	41	369	34.0	584	13	BM328918	BM328918
c	42	368.6	33.9	585	13	BM328644	BM328644
c	43	367.4	33.8	637	10	AM623911	AM623911
c	44	366	33.7	601	10	BE204724	BE204724
c	45	364.2	33.5	585	13	BM499734	BM499734

#### ALIGNMENTS

RESULT 1  
AY108204  
LOCUS AY108204 1542 bp mRNA linear HTC 25-MAY-2002  
DEFINITION Zea mays PCO075514 mRNA sequence.  
ACCESSION AY108204  
VERSION AY108204.1 GI:21211282  
KEYWORDS HTC.  
SOURCE Zea mays.  
ORGANISM Zea mays.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 1542)  
AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,  
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.  
TITLE Zea Mapping Project/DuPont Consensus Sequences for Design of  
Overgo Probes  
JOURNAL Unpublished (2002)  
REFERENCE 2 (bases 1 to 1542)  
AUTHORS Coe,E.C.  
TITLE Direct Submission  
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of  
Missouri, Columbia, MO 65211, USA  
FEATURES  
Location/Qualifiers  
1..1542  
/organism="Zea mays"  
/db\_xref="MaizeDB:634165"  
/db\_xref="taxon:4577"  
/clone="PCO075514"

```
/clone_lib="Maize Mapping Project/DuPont Cornsensus
Library"
/note-"This sequence is part of a project of Est
assemblies resulting from the application of public
contigs to seed Dupont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project."
BASE COUNT      377 a 341 c 408 g 416 t
ORIGIN
Query Match      98.4%; Score 1068.4; DB 11; Length 1542;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1075; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
Qy 1 ATGAAGGCCCTCATCTTCTGGGGGTTTCGGAACCCGCTTCGGCCCTTGTACTCTGAGC 60
Db 202 ATGAAGGCCCTCATCTTCTGGGGGTTTCGGAACCCGCTTCGGCCCTTGTACTCTGAGC 261
Qy 61 TTCCCGAAACCCCTGGTGGATTTTSCAAACAGCCCATCTCTGCACACATCGAAGCT 120
Db 262 TTCCCGAAACCCCTGGTGGATTTTSCAAACAGCCCATCTCTGCACACATCGAAGCT 321
Qy 121 TTGAAGAAGTTGGGGTCACAGAGTGTTTGGTATCACTATCGCCCGAGAGTAATG 180
Db 322 TTGAAGAAGTTGGGGTCACAGAGTGTTTGGTATCACTATCGCCCGAGAGTAATG 381
Qy 181 ATTAATTTCTTCAAGCACTTTTCAGCAATGCTTGGCATCACTATCACTATCGCCCGAG 240
Db 382 ATTAATTTCTTGAAGGACTTTTGAAGATAAATTTGGCATCACTATCACTATCGCCCGAG 441
Qy 241 ACTGAGCCCTTAGGAACCCCTGGCCCTCTGCTCTAGCAAGGCAAGCTTCGGATGGA 300
Db 442 ACTGAGCCCTTAGGAACCCCTGGCCCTCTGCTCTAGCAAGGCAAGCTTCGGATGGA 501
Qy 301 TCTGCGCAGCCATCTTTTGTCTCAACAGTGTATGTCATAGCGAATATCCCATTTCTGAA 360
Db 502 TCTGCGCAGCCATCTTTTGTCTCAACAGTGTATGTCATAGCGAATATCCCATTTCTGAA 561
Qy 361 CTCATCAATTTTCACAACTGTCATGCTGCTGAGGCAACAAATATGGTCACTAAGTGGAT 420
Db 562 CTCATCAATTTTCACAACTGTCATGCTGCTGAGGCAACAAATATGGTCACTAAGTGGAT 621
Qy 421 CAACATCAAAATACGGTGTGCTGTTATGAGGAGGCAACTGCGCAGGCTGCAAGCTTTT 480
Db 622 GAACCACTAATATGGTGTGTTATGAGGAGGCAACTGCGCAGGCTGCAAGCTTTT 681
Qy 481 GTTGAGAAGCCAAATAATTTTGGGTAAACAGATCAATGCTGGGATTTACTTACTGAAC 540
Db 682 GTTGAGAAGCCAAATAATTTTGGGTAAACAGATCAATGCTGGGATTTACTTACTGAAC 741
Qy 541 CCATCTGTCTTGACCGCATTTGAGTGAAGGCAACATCAATTGAGAAGAGTCTTCCT 600
Db 742 CCATCTGTCTTGACCGCATTTGAGTGAAGGCAACATCAATTGAGAAGAGTCTTCCT 801
Qy 601 CAAATTTGAGGTGATCAACAGCTGTCATCAATGCTGCTCAGGCTTTTGGATGATGTT 660
Db 802 CAAATTTGAGGTGATCAACAGCTGTCATCAATGCTGCTCAGGCTTTTGGATGATGTT 861
Qy 661 GGTGAGCCTTAGGACTACATTTACTGCTGCTCTTATCTAGACTCGATAGGAAGAA 720
Db 862 GGTGAGCCTTAGGACTACATTTACTGCTGCTCTTATCTAGACTCGATAGGAAGAA 921
Qy 721 TCAGTGCACAGCTAGCTAGCTGAGGACATGTTTGTGGCATATGCTGCTGATGAGAGC 780
Db 922 TCAGTGCACAGCTAGCTAGCTGAGGACATGTTTGTGGCATATGCTGCTGATGAGAGC 981
Qy 781 GCCAAGATGAGAGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 982 GCCAAGATGAGAGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1041
Qy 841 GTGAGGAGCCGCTGAGGCTTTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 900 GTGAGGAGCCGCTGAGGCTTTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 959
```

```
Db 1042 GTGAGGAGCCGCTGAGGCTTTCCCGCTGCACCTGTCTATCGCGCGCTGCGTATCAAGAAG 1101
Qy 901 CATGCTTGCATCTCAACAGCATATATCGGCTGCGCATCAACTCTTCTGCTAATGGGACGG 960
Db 1102 CATGCTTGCATCTCAACAGCATATATCGGCTGCGCATCAACTCTTCTGCTAATGGGACGG 1161
Qy 961 ATAGAGAAATATGACTATCTTGGGGAGGATGTTTCATGTTGTGTGATGAGGTGTACAGCAAT 1020
Db 1162 ATAGAGAAATATGACTATCTTGGGGAGGATGTTTCATGTTGTGTGATGAGGTGTACAGCAAT 1221
Qy 1021 GCGGCTGTTCTCCACATAAAGAGATCAAGTCAAGCATTTCTGAAGCTTGAGATCGTC 1080
Db 1222 GCGGCTGTTCTCCACATAAAGAGATCAAGTCAAGCATTTCTGAAGCTTGAGATCGTC 1281
Qy 1081 ATGTGA 1086
Db 1282 ATGTGA 1287

RESULT 2
LOCUS      BG321249      980 bp      mRNA      linear      EST 27-FEB-2001
DEFINITION Zm04_04b07_R Zm04_AAFC_EC0RC_cold_stressed_maize_seedlings Zea mays
cDNA clone Zm04_04b07, mRNA sequence.
ACCESSION  BG321249
VERSION     HG321249.1  GI:13150927
KEYWORDS    EST.
SOURCE      Zea mays.
ORGANISM    Zea mays.
REFERENCE   1 (bases 1 to 980)
AUTHORS     Singh,J.A., Wakui,K., Couroux,P., De Moors,A., Harris,L.J., Hattori
            ,J.I., Ouellet,I., Robert,L.S., Sprott,D. and Linker,N.A.
TITLE       Expressed Sequence Tags from Cold-Stressed Maize Seedlings
JOURNAL     Unpublished (2001)
COMMENT     Contact: Singh,J.A.
            Eastern Cereal and Oilseed Research Centre
            Agriculture and Agri-food Canada
            KW Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, K1A
            0C6, Canada
            Tel: (613) 759-1662
            Fax: (613) 759-1701
            Email: singhjadem.agr.ca

FEATURES             Source
     source            1..980
                     /organism="Zea mays"
                     /cultivar="co328"
                     /db_xref="taxon:4577"
                     /clone="Zm04_04b07"
                     /clone_lib="Zm04_AAFC_EC0RC_cold_stressed_maize_seedlings"
                     /tissue_type="leaf, crown"
                     /note="Vector: Bluescript SK-/XhoI-ECORI; Site_1: Eco RI;
                     site_2: Xho I; Lower temperature 50 C / hour from 22 to
                     120C; bring to 50 in 1 hour from 120C. Leave at 50C 2 days
                     photoperiod 16 hours. Light intensity was 125 uE-1.
                     library prepared by in vivo mass excision from amplified
                     library."

BASE COUNT      237 a 225 c 253 g 240 t 25 others
ORIGIN
Query Match      69.4%; Score 753.8; DB 12; Length 980;
Best Local Similarity 96.5%; Pred. No. 2.4e-222;
Matches 769; Conservative 20; Mismatches 6; Indels 2; Gaps 2;
Qy 1 ATGAAGGCCCTCATCTTCTGGGGGTTTCGGAACCCGCTTCGGCCCTTGTACTCTGAGC 60
Db 186 ATGAAGGCCCTCATCTTCTGGGGGTTTCGGAACCCGCTTCGGCCCTTGTACTCTGAGC 243
Qy 61 TTCCCGAAACCCCTGGTGGATTTTSCAAACAGCCCATCTCTGCACACATCGAAGCT 120
Db 246 TTCCCGAAACCCCTGGTGGATTTTSCAAACAGCCCATCTCTGCACACATCGAAGCT 305
```

QY 121 TTGAAGAAGTTGGGTCACAGAGGTGTTTGGCTATCAACTATCGCCAGAGGTAATG 180  
|||||  
Db 306 TTGAAGAAGTTGGGTCACAGAGGTGTTTGGCTATCAACTACCGCCAGAGGTAATG 365  
QY 181 ATTAATTTCTTGAAGACTTTGAGGATAGCTTGGCATCACAATTTACATGCTCCCAAGAG 240  
|||||  
Db 366 ATTAATTTCTTGAAGACTTTGAGGATAGCTTGGCATCACAATTTACATGCTCCCAAGAK 425  
QY 241 ACTGAGCCCTTAGGAACCGCTGGCCCTCTTGCTCTAGCAAGGACAAGCTTCGGGATGA 300  
|||||  
Db 426 ACTRAGCCCTTAGGAACCGCTGGCCCTCTTGCTCTAGCAAGGACAAGCTTCKKATGA 485  
QY 301 TCTGCCAGCCATCTTTTGTCTCAACAGTGTATGATCAATGAAGCAATACCATTTTCTGAA 360  
|||||  
Db 486 TCTGCCAGCCATCTTTTGTCTCAACAGTGTATGATCAATGAAGCAATACBCATTTCTGAA 545  
QY 361 CTCATCAATTTCAAGTGTATGCTGGTGTAGGCAACAATATGCTCACTAAGGTGAT 420  
|||||  
Db 546 CTCATCAATTTCAAGTGTATGCTGGTGTAGGCAACAATATGCTCACTAAGGTGAT 605  
QY 421 GAACCATCAAAATACGGTGTGTGTTATGGAGGAGGCAACTGGCAGGTGGAAAGGTTT 480  
|||||  
Db 606 GAACCATCAAAATACGGTGTGTGTTATGGAGGAGGCAACTGGCAGGTGGAAAGGTTT 665  
QY 481 GTTGAGAGCCAAATAATTTTGGGTAACAAGATCAATGCTGGGATTTACTTACTGAAC 540  
|||||  
Db 666 GTTGAGAGCCAAATAATTTTGGGTAACAAGATCAATGCTGGGATTTACTTACTGAAC 725  
QY 541 CCATCTGTCTTGACCGCATTTAGCTGAGGCCAACAATCAATTTGAGAAAGAGTCTTCCT 600  
|||||  
Db 726 CCATCTGTCTTGACCGCATTTAGCTGAGGCCAACAATCAATTTGAGAAAGAGTCTTCCT 785  
QY 601 CAATTCGAGCTGATCAACAGCTCTATGCAATGGTCTTCAGGTTTGGATGATGTT 660  
|||||  
Db 786 CAATTCGAGCTGATCAACAGCTCTATGCAATGGTCTTCAGGTTTGGATGATGTT 845  
QY 661 GGTGAGCTAGGGACTACATTTACTTGGCTTGGCTTTTATGACTCGATTAGGAAGAAA 720  
|||||  
Db 846 GGTGAGCTAGGGACTACATTTACTTGGCTTGGCTTTTATGACTCGATTAGGAAGAAA 905  
QY 721 TCAGCTGCCAAGCTAGCTAGCTAGGACACATGTTGTCATGCTGCTGTCATGAGAGC 780  
|||||  
Db 906 TCAGCTGCCAAGCTAGCTAGGACACATGTTGTCATGCTGCTGTCATGAGAGC 963  
QY 781 CCCAAGATTGAGAGG 797  
|||||  
Db 964 CCCAAGATTGAGAGG 980  
|||||

RESULT 3  
BG049329  
LOCUS OVI\_18\_E11\_g1\_A002 Ovary 1 (OV1) Sorghum bicolor cDNA, mRNA  
DEFINITION BG049329 694 bp mRNA linear EST 25-JAN-2001  
sequence.  
ACCESSION BG049329  
VERSION BG049329.1 GI:12500932  
KEYWORDS EST.  
SOURCE Sorghum.  
ORGANISM Sorghum bicolor  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Sorghum.  
1 (bases 1 to 694)  
Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt  
,L.H.  
TITLE An EST database from Sorghum: ovaries of varying immature stages  
JOURNAL Unpublished (2000)  
COMMENT Contact: Cordonnier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860

Fax: 706 583 0210  
Email: mmpratt@uga.edu  
Sequences have been trimmed to exclude PolyA, vector and regions  
below Phred quality 16. The threshold for highest quality sequence  
is 20.  
Seq primer: PolyTWix  
High quality sequence start: 36  
High quality sequence stop: 689  
POLYA=No.  
Location/Qualifiers  
1..694  
/organism="Sorghum bicolor"  
/db\_xref="taxon:4558"  
/clone\_lib="Ovary 1 (OV1)"  
/note="Organ: Mix of ovaries of varying immature stages  
from 8-week-old plants; Vector: pBluescript II from lambda  
zap II; Site\_1: XhoI; Site\_2: EcoRI; The library was made  
from poly-A RNA in the cloning vector lambda ZAP II.  
Clones to be sequenced were prepared by mass excision."  
BASE COUNT 169 a 145 c 181 g 199 t  
ORIGIN

Query Match 49.6%; Score 538.2; DB 12; Length 694;  
Best Local Similarity 96.8%; Pred. No. 1.7e-155;  
Matches 549; Conservative 0; Mismatches 18; Indels 0; Gaps 0;  
QY 520 GCTGGGATTTACTTACTGAACCCATCTGCTTTCACCGCATTCAGCTGAGGCCAACATCA 579  
|||||  
Db 1 GCTGGGATTTACTTACTGAACCCATCTGCTTTCACCGCATTCAGCTGAGGCCAACATCA 60  
QY 580 ATTGAGAAAGAGGTCTTCCCTCAAAATTGCAGCTGATCAACAGCTCTATGCAATGGTCTT 639  
|||||  
Db 61 ATTGAGAAAGAGGTCTTCCCTCAAAATTGCAGCTGATCAGAGCTCTATGCAATGGTCTT 120  
QY 640 CCAGGTTTTGGATGGATGTTGGTCAGCCTAGGAGTACATTTACTGGCTTCGCTTTAT 699  
Db 121 CCAGGTTTTGGATGGATGTTGGTCAGCCTAGGAGTACATTTACTGGCTTCGCTTTAT 180  
QY 700 CTAGACTCATTTAGGAAGAAATCAGCTGCCAAGCTAGCTACTGAGGACACATGTTGTTGC 759  
|||||  
Db 181 CTAGACTCATTTAGGAAGAAATCAGCTGCCAAGCTAGCTACTGAGGACACATGTTGTTGC 240  
QY 760 AATGTCTGCTGTCATGAGAGCGCAAGATTGGAGAGGTTGCTGATTGCTGCTGATGTC 819  
Db 241 AATGTCTGCTGTCATGAGAGCGCAAGATTGGAGAGGTTGCTGATTGCTGCTGATGTC 300  
QY 820 GCCATTGGACCTGGGTGTGTGTGGAGGACGGCGTGGAGGCTTTCCCGCTGCTGCTGATG 879  
Db 301 GCCATTGGACCTGGGTGTGTGTGGAGGATGGCGTGGAGGCTTTCCCGCTGCTGCTGATG 360  
QY 880 CGCGGGTCCGTATCAAGAAGCATGCTTGCATCTCAACAGCATTTATCGGCTGGCACTCA 939  
|||||  
Db 361 CGTGTGTGCGCATCAAGAAGCATGCTTGCATCTCAACAGCATTTATCGGCTGGCACTCA 420  
QY 940 ACTGTTGGTCAATGGCACGATAGAGATATCATTCTCTGGGAGGATGTTTCATGTC 999  
Db 421 ACTGTTGGTCAATGGCACGATAGAGATATCATTCTCTGGGAGGATGTTTCATGTC 480  
QY 1000 TGTGATGAGGTGTACAGCAATGGCGTGTGTTCTCCACATAAAGAGATCAAGTCAAGC 1059  
Db 481 TGTGATGAGGTGTACAGCAATGGCGTGTGTTCTCCACATAAAGAGATCAAGTCAAGC 540  
QY 1060 ATTCTGAAGCCCTGAGATCGTCATGTGA 1086  
Db 541 ATTCTGAAGCCCTGAGATCGTCATGTGA 567  
|||||

RESULT 4  
BE455732  
LOCUS BE455732  
DEFINITION BE455732 836 bp mRNA linear EST 22-OCT-2001  
HVSMEG0019G12f Hordeum vulgare pre-anthesis spike EST library  
HVCNDA0008 (white to yellow anther) Hordeum vulgare cDNA clone  
HVSMEG0019G12f, mRNA sequence.



```

ACCESSION      BE455732
VERSION        BE455732.2  GI:13155059
KEYWORDS
SOURCE         Hordeum vulgare.
               Hordeum vulgare.
               Hordeum vulgare.
ORIGIN         Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
               ; Triticeae; Hordeum.
               1 (bases 1 to 836)
REFERENCE      Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Begum, D., Frisch, D., Fenton
AUTHORS        Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi, D.W., Fenton
               , R.D., Close, S.J., Oates, R. and Main, D.
TITLE          Development of a genetically and physically anchored EST resource
               for barley genomics: Morex pre-anthesis spike cDNA library
               Unpublished (2001)
JOURNAL        On Jul 26, 2000 this sequence version replaced gi:9465899.
COMMENT        Contact: Wing RA
               Clemson University Genomics Institute
               Clemson University
               100 Jordan Hall, Clemson, SC 29634, USA
               Tel: 864 656 7288
               Fax: 864 656 4293
               Email: rwing@clemson.edu
               Total hq bases = 404
               Seq primer: AATTAACCTCTCACTAAAGGG
               high quality sequence stop: 563.
               Location/Qualifiers
               1..836
               /organism="Hordeum vulgare"
               /cultivar="Morex"
               /db_xref="taxon:4513"
               /clone="HVMSEG0019G12#n"
               /clone_lib="Hordeum vulgare pre-anthesis spike EST library"
               /HVCIDNA0008 (white to yellow anther)
               /tissue_type="pre-anthesis spike"
               /lab_host="SOLR"
               /notes="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI;
               Plants were grown in the greenhouse at the University of
               California, Riverside (Fenton, SJ Close, TJ Close). Whole
               spike with awns trimmed were collected at white, green and
               yellow anther stages (Fenton). Total RNA was prepared from
               each pool, equal quantities of all three RNA pools were
               combined, poly(A) RNA was purified from the mixture, one
               primary unamplified cDNA library was made, and 1 million
               pfu were in vivo excised to give pBluescript SK(-) cDNA
               phagemids. These steps were performed in the TJ Close lab
               (Choi) at the University of California, Riverside.
               Phagemids were plated and picked at the Clemson University
               Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins
               and Wing) Plasmid DNA preparations, DNA sequencing and
               sequence analysis were performed at CUGI (Wing, Yu, Frisch
               , Henry, Simmons, Oates, Rambo, Main). The sequence has
               been trimmed to remove vector sequence and contains a
               minimum of 100 bases of phred value 20 or above. For more
               details on library preparation and sequence analysis see
               http://www.genome.clemson.edu/projects/harley TO order
               this clone see http://www.genome.clemson.edu/orders Also
               see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
               Genetically and physically anchored EST resources for
               barley genomics. Barley Genetics Newsletter 31:29-30.
               (http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"
               205 a 216 c 218 g 197 t
BASE COUNT
ORIGIN
Query Match 45.2%; Score 490.4; DB 10; Length 836;
Best Local Similarity 82.3%; Pred. No. 1.4e-140;
Matches 611; Conservative 0; Mismatches 126; Indels 5; Gaps 4;

QY 1 ATCAAGCCCATCATTCCTCCGGGGTTTCGGACACCGCCCTTCGGCCTTCACCTCTCCACC 60
Db 80 ATGAAGCGCTCATTCCTCGTCGGGGGGCTTCGGGACCGCCCTTCGGCGCTCTCACGCTCAGC 139
QY 61 TTCCCGAAACCCCTCGCTGGGATTTTGCAACCAAGCCCATGATTCGTGCACCAAGATCGAAGCT 120

```

Email: mmpratt@uga.edu

Sequences have been trimmed to exclude PolyA, vector, and regions below Phred quality 16. The threshold for highest quality sequence is 20. Three-prime sequences, which are obtained with PolyTmix or T7 sequencing primer, are presented as the reverse complement.

Seq primer: JEN REV

High quality sequence stop: 461

POLYA-No.

#### FEATURES

source

Location/Qualifiers  
1. .524  
/organism="Sorghum bicolor"  
/cultivar="B7x623"  
/db\_xref="taxon:4558"  
/clone\_lib="Pathogen-infected compatible 1 (PIC1)"  
/tissue\_type="Leaves"  
/dev\_stage="4-week-old seedlings infected with Colletotrichum graminicola"  
/note="vector: pbluescript II SK(-) from Lambda Zap II; Site1: XhoI; Site2: EcoRI; Four-week-old sorghum seedlings were sprayed with spore suspension prepared from 3-week-old FRM421, a sorghum isolate of the anthracnose pathogen Colletotrichum graminicola. Inoculated plants were kept in a 25 C dark growth chamber with 100% relative humidity for 24 hr, followed by 12/12 hr of light/dark cycle at 25 C with 90% relative humidity for another 24 hr. All leaves were harvested and quick frozen with liquid nitrogen and stored in a -80 C freezer. The library was made from poly-A RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision.  
WARNING: While most or all ESTs are expected to derive from the host plant, no effort was made to eliminate ESTs deriving from the pathogen."

BASE COUNT 138 a 104 c 140 g 142 t

ORIGIN

Query Match 44.8%; Score 486; DB 13; Length 524;

Best Local Similarity 96.1%; Pred. No. 2.2e-139;

Matches 498; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 259 GCTGGCCTCTGTCTAGCAAGGACAAAGCTTGGGATGATCGGCCAGCATCTTT 318

Db 7 GGTGGTCTCTGTCTAGCAAGGACAAAGCTTGGGATGATCGGCCAGCATCTTT 66

QY 319 GTCTCAACAGTGTGTCATAGGCAATACCATTTGCTGAACATCATCAAAATTCACAA 378

Db 67 GTCTCAACAGTGTGTCATAGGCAATACCATTTGCTGAACATCATCAAAATTCACAA 126

QY 379 TGTCTATGGTGTGAGGCAACAATATGCTACTAAGTGGATGAACCATCAAAATACGGT 438

Db 127 TGTCTATGGTGTGAGGCAACAATATGCTACTAAGTGGATGAACCATCAAAATATGCT 186

QY 439 GTTCTGTTATGGAGGACCAACTGGCAGGCTGGNAAGTTGTTGAGAAGCCAAATA 498

Db 187 GTTCTGTTATGGAGGACCAACTGGCAGGCTGGNAAGTTGTTGAGAAGCCAAATA 246

QY 499 TTTCTGGTTAAAGATCAATGCTGGGATTTACTTACTGAACCCATCTGCTTACCCG 558

Db 247 TTTCTGGTTAAAGATCAATGCTGGGATTTACTTACTGAACCCATCTGCTTACCCG 306

QY 559 ATTGAGCTGAGGCCAACATCAATTGAGAAGAGTCTTCCCTCAAATTCAGCTGATCAA 618

Db 307 ATTGAGCTGAGGCCAACATCAATTGAGAAGAGTCTTCCCTCAAATTCAGCTGATCAG 366

QY 619 CAGCTCTATGCAATGGCTTCCAGGTTTTGGATGGATTTGGTCAGCTTAGGACATAC 678

Db 367 AAGCTCTATGCAATGGCTTCCAGGTTTTGGATGGATTTGGTCAGCTTAGGACATAC 426

QY 679 ATTACTGGCTTGCCTTTTACTAGACTCGATTAGGAAGAATCAGTGCCTGCAAGTACGT 738

Db 427 ATTACTGGCTTGCCTTTTACTAGACTCGATTAGGAAGAATCAGTGCCTGCAAGTACGT 486

QY 739 ACTGGAGCACATGTTGTTGGCAATGTGCTGGTCATGA 776

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 487 ACTGGAAACACATGTTGTCGAATGCTGTCATGA 524

RESULT 6

BE405374

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (2000)

Contact: Olin Anderson

US Department of Agriculture, Agriculture Research Service, Pacific

West Area, Western Regional Research Center

800 Buchanan Street, Albany, CA 94710, USA

Tel: 5105595773

Fax: 5105595818

Email: oanderson@pw.usda.gov

Sequence have been trimmed to remove vector sequence and low

quality sequence with phred score less than 20

Seq primer: Strategene SK primer.

Location/Qualifiers

1. .597

/organism="Triticum aestivum"

/cultivar="Chinese Spring"

/db\_xref="taxon:4565"

/clone="WHE1213\_F04\_K07"

/clone\_lib="Wheat etiolated seedling root cDNA library"

/tissue\_type="Root"

/dev\_stage="Five day old etiolated seedling"

/lab\_host="E. coli SOLR"

/note="Vector: Lambda Uni-ZAP XR, excised phagemid;

Site\_1: EcoRI; Site\_2: XhoI; Seeds were surface-sterilized

, germinated and grown aseptically in the dark at room

temperature on filter paper with water, nystatin and

cefotaxime in covered crystallization dishes. Roots were

harvested. The tissue, total RNA, and poly(A) RNA were

prepared, a cDNA library was made, and the cDNA clones

were in vivo excised to give pBluescript phagemids in the

to Close lab (Choi, Close, Fenton) at the University of

California, Riverside. Plasmid DNA preparations and DNA

sequencing were performed in the OD Anderson lab (all

other authors)."

BASE COUNT 161 a 138 c 144 g 154 t

ORIGIN

Query Match 44.5%; Score 483.4; DB 10; Length 597;

Best Local Similarity 88.1%; Pred. No. 1.6e-138;

Matches 526; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 37 CGCCTTCGGCCTTGTACTCTGAGCTTCCCGAAACCCCTCGTGATTTTGCACAAAGCCC 96

Db 1 CGCCTTCGGCCTTGTACTCTGAGCTTCCCGAAACCCCTCGTGATTTTGCACAAAGCCC 60

QY 97 ATGATTCTGCACAGATCGAAGCTTTGAAAGAGTTGGGCTGCAGAGGTGGTTTTCGCT 156

Db 61 ATGATTCTGCACAGATCGAAGCTTTGAAAGAGTTGGGCTGCAGAGGTGGTTTTCGCT 120

QY 157 ATCAACTATCGCCAGAGGTAAATGATTAATTTCTTGAAGGACTTTGAGGATAAGCTTGGC 216

Db 121 ATCAACTATCGCCAGAGGTAAATGATTAATTTCTTGAAGGACTTTGAGGATAAGCTTGGC 180

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||







1 (bases 1 to 651)  
Sato,K., Saisho,D. and Takeda,K.  
Barley EST sequencing project in NIG and Okayama Univ  
Unpublished (2002)  
Contact: Tadasu Shin-1  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.

FEATURES  
source  
    ..651  
        /organism="Hordeum vulgare subsp. vulgare"  
        /cultivar="Haruna Niho"  
        /db\_xref="taxon:112509"  
        /clone="baal6n22"  
        /clone\_lib="K. Sato unpublished cDNA library, cv. Haruna  
            Niho adult, heading stage top three leaves"  
        /tissue\_type="top three leaves"  
        /dev\_stage="adult, heading stage"

BASE COUNT     170 a   182 c   145 g   154 t

ORIGIN

Query Match       41.2%; Score 447.4; DB 10; Length 651;  
Best Local Similarity 86.0%; Pred. No. 2.6e-127;  
Matches 496; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

Qy 510 CAAGATCAATGCTGGGATTACTTACTGAACCCATCTGTCTTGACCGCATTTGAGCTGAG 569  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 651 CAAGATCAATGCAGAAATTACTTGTAAATCCATCTGTCTTGACCGCATTCGAATTGAA 592  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Qy 570 GCCAACATCAATTGAGAAAGAGGCTTCCCCTCAAATTGCAGCTGATCAACAGCTCTATGC 629  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 591 GCCAACTTCANTCAGAAGGAGGCTTCCCCTCGAATGTCAGCTGATCAGAGCTCTACGC 532  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Qy 630 AATGCTCTCCAGGTTTTTGGATGGATGTTGGTCAGGCTTAGGAGCTACATACTGGGCTT 689  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 531 CATGGTCTCTCCAGGCTTTTGGATGGATATTGGCCAGGCCAAGGGATTACATCACTGGGCTT 472  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Qy 690 CGGTCTTTATCTAGACTCGATTAGGAAGAAATACGCTGCCAAGCTAGCTACTGGAGCAC 749  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 471 CGGTCTTATCTAGATTACATTAGGAAGAAATACGCTGCCAAGCTGGCCGTCGGAGAAC 412  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Qy 750 TGTTGTTGGCAATGTGCTGGTCATGAGAGCGCCAAGATTGGAGAAGTTCTCTGATTGG 809  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 411 TTGTTGTTGGGAACGTCCTGGTGCAGAGATGCCAAGATTGGGGAGGCTCTCTGATTGG 352  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Qy 810 TCCTGATGTCGCCATTTGACCCTGGGTGTGTTGTGGAGAGCGGCTGAGGCTTTCCCGCTG 869  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 351 CCCGTGATGTTGCTATTGGACCTGGATGCTGTTGGAGGATGTTGTGAGGCTCTCCCGATG 292  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Qy 870 CACTGTCAATGCGCGGCTGGCTATCAAGAACATGCTTGCATCTCAACACGATATTCGG 929  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 291 CACGGTGATGGTGGTGTACCGCTTAAGAACATGCGGTGCATCTCAACACGATCATCGG 232  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Qy 930 CTGGCACTCAACTGTTGGTCAATGGGACGATAGAGAATATGACTATCTCTGGGGAGGA 989  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 231 CTGGCATTCGACGTCGGGCAATGGGACGATAGAGNATATGACTATCTCTGGGAGAGA 172  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Qy 990 TGTTCATGTGTGATGAGGTGTACAGCAATGGCGGTGTGTTCTTCCACATAAGAGAT 1049  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 171 TGTACATGTTGGTGACGAGGTCTACAGCAACGCGGTGTTGTTCTTCCGCGACAAGGAGAT 112  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Qy 1050 CAAGTCAAGCATTTCTGAAGCCTGAGATCGTCAATGTGA 1086  
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 111 CAAGTCAAGCATCTTAAGCCTGAGATCGTCAATGTGA 75  
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12  
BE844694       BR844694       794 bp   mRNA       linear       EST 17-MAY-2001  
LOCUS       AD018017 AD A. thaliana (Col-0 ql1) library enriched for  
DEFINITION

```

QY 349 CGATTGTGTAAGTCAATCAATTTACAGAGTGTGAGTGGTGGAGGACAAATTAATGGTC 408
DB 320 CCTCTAAAGAAATGCTTACAGTTCACAAATCTCAGGTCGGAGAGCCCTCCATATGGGTA 379
QY 409 ACTAGGTGGATCAACCAATCAATACAGTGTGTGGTTATGAGGAGGACACTGGCAGG 468
DB 380 ACAGAGGTGGATGACCGTGGAAATATGGAGTGGTTTATGGAGAGAGCACTGGGAAGA 439
QY 469 GTGGAAGGTTGTGACAGAGCCAAATATTTCTGGGTAACAGATCAATGTCGCAAT 528
DB 440 GTGGAGAGTGTGCGAAGAGCCAAACTGTATGAGTGAACAGATCAACGCTGGGAT 499
QY 529 TACTTACTGAACCCATCTGCTGACCGCATTTAGCTGAGCGCCAAACATCAATGAGAAA 588
DB 500 TATCTTCTGAACCCATCTGCTTCTGATGAATGAGTGAAGACCGACTTTCATCCAAA 559
QY 589 GAGGTCTTCCCTCAATTTGACAGTCAACAGCTCTATGCAATGGTCTTCCAGGTTTT 648
DB 560 GAGACTTTTCTTAAGATTCGACAGCGGAGGCGCTCTATGCTATGCTGTACAGGGTTT 619
QY 649 TGGATGGATGTCGTCAGCTGAGGACATACATGCTGCTTCCCTCTTATCTAGACTCC 708
DB 620 TGGATGGATGTCGTCAGCTGAGGACATACATGCTGCTTCCCTCTTATCTAGACTCC 679
QY 709 ATTAGAGAAATCACTGCGCAAGCTAGCTACTGAGGACATGCTGCTGCAATGTGCTG 768
DB 680 CTAGAGAAATCTCTGCGCAA-ATACCACTGCGCCACACATAGTGGGAATTCCTCT 738
QY 769 GTGCATGAGAGCGCAAGATGGAGAGGTGTCTGATTTGGTCTGATGT 818
DB 739 TGTGACGAACCGCTACAAATGGGGAAGGATGTTTGTATGGCCACAGACGT 788

RESULT 13
BQ806351
LOCUS
DEFINITION
WHE3577_G06_N112s wheat developing grains cdna library Triticum
aestivum CDNA CLONE WHE3577_G06_N11, mRNA sequence.
ACCESSION
BQ806351
VERSION
BQ806351.1 GI:22030560
KEYWORDS
EST.
SOURCE
bread wheat.
ORGANISM
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
1 (bases 1 to 619)
Altenbach, S., Anderson, O.D., Chao, S., Chin, A., Close, T.J., Cronin
, K., Crossman, C., Penton, R.D., Lazo, G.R., Pham, J., Rausch, C.J.,
Wilson, C. and Woo, J.
The structure and function of the expressed portion of the wheat
genomes - Developing grains cdna library
Unpublished (2002)
Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: Oanderson@pw.usda.gov
Sequences have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: SK primer.
Location/Qualifiers
1. .619
/organism="Triticum aestivum"
/cultivar="Butte 86"
/db_xref="taxon:4565"
/clone="WHE3577_G06_N11"
/clone_lib="Wheat developing grains cdna library"
/tissue_type="whole grains"
/dev_stage="3-44 days post anthesis seed"

```

```

/lab_host="E. coli SOLR"
/Note="Vector: Lambda ZAP II, excised phagemid; Site 1:
EORI; Plants were grown under six following different
environmental regimes in greenhouse, Environment 1)
24oc/17oc day/night, well-watered, with post-anthesis
fertilizer, Environment 2) 24oc/17oc day/night,
well-watered, without post-anthesis fertilizer,
Environment 3) 37oc/17oc day/night, well-watered, with
post-anthesis fertilizer, Environment 4) 37oc/17oc
day/night, well-watered, without post-anthesis fertilizer,
Environment 5) 37oc/17oc day/night plus drought, with
post-anthesis fertilizer, Environment 6) 37oc/17oc
day/night plus drought, without post-anthesis fertilizer,
developing wheat grains from the following were excised
and frozen in liquid nitrogen, Environment 1 at 3, 5, 7,
8, 10, 12, 16, 20, 24, 28, 32, 36, 40, 44 DPA Environment
2 at 3, 5, 7, 8, 10, 12, 16, 20, 24, 28, 32, 36, 40, 44,
DPA Environment 3 at 3, 5, 7, 8, 10, 12, 16, 20, 24, 28,
32, 34 DPA Environment 4 at 3, 5, 7, 8, 10, 12, 16, 20, 24
, 28, 32, 34 DPA Environment 5 at 3, 5, 7, 8, 10, 12, 16,
20, 24, 28, 30 DPA Environment 6 at 3, 5, 7, 8, 10, 12, 16
, 20, 24, 28, 30 DPA and total RNA was prepared by S.
Altenbach and K. Cronin at USDA-ARS, Albany, CA. A cDNA
library was made using poly (A) RNA, and the cDNA clones
were in vivo excised to give pBluescript SK(-) phagemids
in the TJ Close lab (Chin, Close, Penton) at the
University of California, Riverside. Plasmid DNA
preparations and DNA sequencing were performed in the OD
Anderson lab (others).
BASE COUNT 147 a 140 c 181 g 150 t 1 others
ORIGIN
Query Match 40.3%; Score 438; DB 14; Length 619;
Best Local Similarity 85.3%; Pred. NO. 2.1e-124;
Matches 511; Conservative 0; Mismatches 86; Indels 2; Gaps 2;
QY 489 GCCAAAATATTTGGGTAAAGATCAATGCTGGGATTTACTTACTGAACCATCTGT 548
DB 3 GGCACGAGGATTTGTAGGCAACAAGATCAATGCTGGAATTTACTCTTAAATCCCTCTGT 62
QY 549 CQTFGACCGCATTTGAGCTGAGCCACATCAATTTGAGAAGAGGCTCTTCCCTCAAATTC 608
DB 63 CCTGGACCGCATCGAGTTAAAGCCAACTTCAATCGAGAAAGAGGCTTTCCGCAATTC 122
QY 609 AGCTGATCAACACTCTATGCAATGCTCCATGCTTCCAGGTTTGGATGCTGTGTCACCC 668
DB 123 TGTGATCAACACTCTACGCCATGCTTCCAGGTTTGGATGATATTTGCCAGCC 182
QY 669 TAGGGACTACATTACTGGCTTGGCTTTATCTAGACTCGATTAGGAAGAAATCAGCTGC 728
DB 183 AAGGACTACATTACTGCGCTTGGCTTTATCTAGACTCGATTAGGAAGAAATCAGCTGC 242
QY 729 CAAGCTAGCTACCTGGAGCACATCTTGTGCAATGTGCTGTCATGAGAGCGCCAGAT 788
DB 243 CAAGCTGGCCGCGGAGCACATATTGGGGAAATGCTCTGTCGACGACGCGCC-AGAT 301
QY 789 TGCAGAGGTTGCTGATGCTGCTGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCG 848
DB 302 TGGGGAGGGTTGCTGATTTGCTGATTTGCTGATTTGCTGATTTGCTGATTTGCTGATTT 361
QY 849 CGGCGTGGAGCTTTCGCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 908
DB 362 TGTGTCGAGCTCTCCGATGCGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 421
QY 909 CATCTCAAAACAGCATTTATCGGCTGGCACTCACTGTTGGTCAATGGCCAGGATAGAGAA 968
DB 422 CATCTCAAAACAGCATTTATCGGCTGGCACTCACTGTTGGTCAATGGCCAGGATAGAGAA 481
QY 969 TATGACTATCTCGGGGAGGATGTTTCATGT-GTGTGATGAGGTGTACAGCAATCGCGGTG 1027
DB 482 TATGAGCATCTCGGGGAGGATGTACATGTAGNGTGTACAGAGGTCTACACAATCGCGGTG 541
QY 1028 TTGTTTCCCAACATAAAGACATCAAGTCAAGCATTTCTGACCCCTGACATCTCTCATCTCA 1086

```



```
||||| 542 TTGTTCTCCCGCACAAAGAGATCAAGTCAAGCATCTCTGAAGCCTGAGATCGTCATGTGA 600

RESULT 14
AV928662/c
LOCUS AV928662 K. Sato unpublished cDNA library, cv. Haruna Nijo second
DEFINITION leaf stage seedling leaves Hordeum vulgare subsp. vulgare cDNA
clone basd15d19 3', mRNA sequence.
ACCESSION AV928662
VERSION AV928662.1 GI:18224459
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare.
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Hordeum.
1 (bases 1 to 673)
Sato,K., Saisho,D. and Takeda,K.
Barley EST sequencing project in NIG and Okayama Univ
Unpublished (2002)
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES
source
Location/Qualifiers
1..673
/organism="Hordeum vulgare subsp. vulgare"
/cultivar="Haruna Nijo"
/db_xref="taxon:112509"
/clone="basd15d19"
/clone_lib="K. Sato unpublished cDNA library, cv. Haruna
Nijo second leaf stage seedling leaves"
/tissue_type="seedling leaves"
/dev_stage="second leaf stage"
BASE COUNT 179 a 187 c 147 g 159 t 1 others
ORIGIN

Query Match 40.3%; Score 437.4; DB 10; Length 673;
Best Local Similarity 85.7%; Pred. No. 3.4e-124;
Matches 486; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 520 GCTGGGATTTACTTACTGAACCCATCTGCTTGCACGCAATGAGCTGAGGCCAACATCA 579
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
673 GCAGGAATTTACTTGTAAATCCATCTGTCTTGACCGCATCGAATGAAGCCAACTTCA 614
QY 580 ATTCAGAAAGAGCTTCCCTCAATTCAGCTGATCAACAGCTCTATGCAATGGTCTTT 639
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
613 ATCAGAAAGAGGCTTCCCTCGAATTCAGCTGATCAGAAAGCTCTAGCCATGGTCTCT 554
QY 640 CCAGGTTTTGGATGGATGTTGGTCAGCTAGGACTACATTAATCTGGCTTCGCTTTAT 699
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
553 CCAGGCTTTGGATGGATATGGCCAGCAAGGATTCATCATCTGGCTTCGCTCTAT 494
QY 700 CTAGCTCGATTAGGAAGAATCAGCTGCCAAGCTAGCTACTGGAGCAATGTTGTTGGC 759
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
493 CTAGATCACTTAGGAAGAAATCAGCTGCCAAGCTGGCCGTCGAGAACATTTGTTGGG 434
QY 760 AATGTCGTGTCATGAGAGCGCCAAAGATTGGAGAAGTTGTCGTGATGGTCCGTGATGC 819
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
433 AACGTCCTGTCACGAGAGTGCCAAAGATTGGGGAGGCTGTCTGATTTGGCCCTGATGT 374
QY 820 GCATTGGACCTGGTCTGTTGGAGAGCGCGTGAGGCTTTCCCGCTGCACCTGTCATG 879
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
373 GCTATTGGACCTGGATGCGTGTGGAGGATGGTGTGAGGCTCTCCCGATCACCGGTGAT 314
QY 880 CGCGGGCTGCTATCAAGAAGCATGCTTGATCTCAACAGCATATATCGCTGGCACATCA 939
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
313 CGTGGGTAGCCATTAGAGAGCATGCGTGATCTCAACAGCATATCGCGCTGGCAITCG 254
```

```
QY 940 ACTGTTGGTCAATGGCACGGATAGAAATATGACTATCTCTGGGGAGGATGTTCTATGTG 999
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
253 ACGGTCGGGCAATGGCACCATAGAGATATGACTATCTCTGGGAGAGATGTACATGTT 194
QY 1000 TGTGATGAGTGTACAGCAATGGCGTGTGTTCTCCCATATAAGAGATCAAGTCAAGC 1059
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
193 GGTGACGAGTGTACAGCAACGCGGTGTGTTCTCCGACAAAGAGATCAAGTCAAGC 134
QY 1060 ATTCTGAAGCCTGAGATCGTCATGTGA 1086
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
133 ATCCTTAAGCCTGAGATCGTCATGTGA 107

RESULT 15
BG414225
LOCUS BG414225
DEFINITION HVSMEK0002B16f Hordeum vulgare testa/pericarp EST library
sequence.
ACCESSION BG414225
VERSION BG414225.2 GI:16328388
KEYWORDS EST.
SOURCE Hordeum vulgare.
ORGANISM Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Hordeum.
1 (bases 1 to 808)
Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Kannangara,G., von
Wetstein,D., Akhunov,E., Chin,A., Choi,D.W., Fenton,R.D., Kianian
,P., Otto,C., Simons,K., Zhang,D., Begum,D., Frisch,D., Yu,Y.,
Henry,D., Palmer,M., Rambo,T., Simons,J., Oates,R. and Main,D.
Development of a genetically and physically anchored EST resource
for barley genomics: Morex testa/pericarp cDNA library
Unpublished (2001)
On Mar 13, 2001 this sequence version replaced gi:13319872.
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total hg bases = 515
Seq primer: AATTAACTCTCTACTAAAGG
High quality sequence stop: 721.
Location/Qualifiers
1..808
/organism="Hordeum vulgare"
/cultivar="Morex"
/db_xref="taxon:4513"
/clone="HVSMEK0002B16f"
HVCDNA0013 (normal)
HVCDNA0013 (normal)
/tissue_type="testa/pericarp"
/lab_host="TJClz1"
/Note="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI;
Plants were raised from seeds in a Controlled Environments
growth chamber maintained in continuous light at 180C, and
testa and pericarp were dissected from developing kernels
at Washington State University, Pullman, WA (Kannangara,
von Wetstein). Total RNA was prepared, poly(A) RNA was
purified, one cDNA library was made, and 1 million pfu
were in vivo excised to give pBluescript SK(-) cDNA
phagemids in the TJ Close lab at the University of
California, Riverside (Akhunov, Chin, Choi, Close, Fenton,
Kianian, Otto, Simons, Zhang). Phagemids were plated and
picked at the Clemson University Genomics Institute (CUGI)
(Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA
preparations, DNA sequencing and sequence analysis were
performed at CUGI (Wing, Yu, Frisch, Henry, Simons, Oates
, Rambo, Main). The sequence has been trimmed to remove
```



vector sequence and contains a minimum of 100 bases of  
phred value 20 or above. For more details on library  
preparation and sequence analysis see  
<http://www.genome.clemson.edu/projects/barley>. To order  
this clone see <http://www.genome.clemson.edu/orders> Also  
see Close TJ, Wing R, Kleinjofcs A, Wise R (2001)  
Genetically and physically anchored EST resources for  
barley genomics. Barley Genetics Newsletter 31:29-30.  
(<http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html>)"

BASE COUNT	171 a	227 c	255 g	155 t	
ORIGIN					
Query Match	40.1%	Score 435;	DB 12;	Length 808;	
Best Local Similarity	75.5%	Pred. No. 2.2e-123;			
Matches 540;	Conservative 0;	Mismatches 175;	Indels 0;	Gaps 0;	
QY	1	ATGAAGGCCCTCATCTTGTGGGGGTTTCGGAAACCCGCTTCGGCCTTTGACTCTGACC	60		
Db					
QY	59	ATGAAGGCCCTCATCTTGTGGGGGTTTCGGAAACCCGCTTCGGCCTTCGCTCAGC	118		
Db					
QY	61	TTCCCGAAACCCCTCGTGGATTTCGAAACAGCCCATGATTCGACCCAGATCGAAGCT	120		
Db					
QY	119	GTGCCCAACCCCTCGTGGATTTCGCAACAGCCCATGATTCGACCCAGATCGAAGCT	178		
Db					
QY	121	TTGAAGAAGTTGGGTACAGAGTGGTTTGGCTATCAACTATGCCCCAGAGGTAATG	180		
Db					
QY	179	CTGAAGATGTGGAGTCACAGAAAGTTGTCTGGCGATCAATTACGACGAGAGTCAAG	238		
Db					
QY	181	ATTAATTTCTCAAGACATTTGAGATTAAGCTTGGCATCAACAAATACATCTCCCAAGAC	240		
Db					
QY	239	CTCAACTTTCTCANGACTTTCGAGAGCAAGCTTGGCATCAAGATCACCTTGTCCCAAGGAG	298		
Db					
QY	241	ACTGAGCCCTTAGGAACCGCTGGCCCTCTGTCTAGCAAGGCAAGCTTCGGATGGA	300		
Db					
QY	299	ACAGACCAATGGAACCCCGGACCGCTGGCCCGCCGACNAAGCTCAGCAGCGA	358		
Db					
QY	301	TCTGCCAGCCATTCTTTGTCTCAACAGTGTGTATAGGGAATACCCATTTCGTGAA	360		
Db					
QY	359	TCCGGGAGCCCTTCTTGTCTCAACAGTGTGTATAGGGAATACCCATTTCGTGAA	418		
Db					
QY	361	CTCATCAAAATTCACAGTGTGTGTGTAGGGAATACCCATTTCGTGAA	420		
Db					
QY	419	CTCATCGAGTTCCACAGTCCCATGGCGGAGGCGACGATCATGGTACCAAGGTGGAC	478		
Db					
QY	421	GAACCATCAAAATTCACAGTGTGTGTGTAGGGAATACCCATTTCGTGAA	480		
Db					
QY	479	CAGCCTTCAGTACCGCTGTGTGTGTAGGGAATACCCATTTCGTGAA	538		
Db					
QY	481	GTGAGAACCCAAAATATTTGTGGTAAACAGATCAATGTGGATTACTTACTGAAC	540		
Db					
QY	539	GTGGAGAAGCCCAAGGTGTCTGTGGCAACAGATCAACCCCGGATCTACCTGCTGAC	598		
Db					
QY	541	CAATCTGTCTGACCCGATGAGTGTGGGCAACATCAATTTGAGAAAGGCTTTCCT	600		
Db					
QY	599	CCGTGCTGTGGACCGCATGAGCTGAAGCCGACGTCCATCGAAGAGGAGGTTCCTCG	658		
Db					
QY	601	CAATCTGTCTGACCCGATGAGTGTGGGCAACATCAATTTGAGAAAGGCTTTCCT	660		
Db					
QY	659	CGTATCGCCCGGCAAAAGGCTGTTCGCCATGCTGCTGCTGCTGCTGCTGCTGCTG	718		
Db					
QY	661	GTCTAGGCTAGGAGCTACATTACTGCTTGTGCTTTATCTATAGACTCGATTAGGA	715		
Db					
QY	719	CGGACCGCAAGAGACTTACTTACCGGCTGAGGCTCTACCTGCTGCTGCTGCTGCT	773		
Db					

Search completed: November 27, 2002, 04:03:48  
Job time : 1994 SECS

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 04:05:27 ; Search time 67 Seconds  
(without alignments)  
717.962 Million cell updates/sec

Title: us-09-374-967-2

Perfect score: 1864

Sequence: 1 MKALILVGGFGRRLRLTL.....GVLPKHKSSILKPEIVM 361

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_101002.\*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1835	98.4	361	21	Arabidopsis thalia
2	1681	90.2	361	21	Arabidopsis thalia
3	1681	90.2	361	21	Arabidopsis thalia
4	1681	90.2	361	21	Arabidopsis thalia
5	1680	90.1	361	20	A. thaliana enviro
6	1671	89.6	329	21	Potato GDP-mannose
7	1540	82.6	302	21	Arabidopsis thalia
8	1517	81.4	329	21	Arabidopsis thalia
9	1517	81.4	329	21	Arabidopsis thalia
10	1402	75.2	302	21	Arabidopsis thalia

11	1402	75.2	302	21	Arabidopsis thalia
12	1183.5	63.5	360	21	Human ORFX ORF1521
13	1183.5	63.5	360	21	Human transferase
14	1175	63.0	363	22	NDP-hexose pyropho
15	1173	62.9	369	22	Drosophila melanog
16	1173	62.9	369	22	Drosophila melanog
17	1160	62.2	387	22	Human protein sequ
18	682	36.6	152	21	Zea mays protein f
19	669	35.9	143	21	Zea mays protein f
20	631.5	33.9	240	22	Novel human enzyme
21	592.5	31.8	420	22	Human protein sequ
22	592.5	31.8	420	22	Human protein sequ
23	554	29.7	411	21	Arabidopsis thalia
24	554	29.7	415	21	Arabidopsis thalia
25	550	29.5	411	21	Arabidopsis thalia
26	550	29.5	415	21	Arabidopsis thalia
27	545	29.2	402	21	Arabidopsis thalia
28	541	29.0	402	21	Arabidopsis thalia
29	529.5	28.4	415	21	Arabidopsis thalia
30	529.5	28.4	431	21	Arabidopsis thalia
31	521	28.0	448	22	Drosophila melanog
32	520.5	27.9	403	21	Zea mays protein f
33	512.5	27.5	111	21	Zea mays protein f
34	493.5	26.5	385	21	Arabidopsis thalia
35	488	26.2	359	22	Mycobacterium tube
36	484.5	26.0	375	21	Arabidopsis thalia
37	478	25.6	416	22	Putative nucleosid
38	465	24.9	92	21	Zea mays protein f
39	465	24.9	152	21	Zea mays protein f
40	444	23.8	361	22	Putative P. abyssi
41	442.5	23.7	362	22	Corynebacterium gl
42	440	23.6	340	21	Arabidopsis thalia
43	431.5	23.1	348	22	C. glutamicum prote
44	411.5	22.1	346	21	Arabidopsis thalia
45	409	21.9	306	21	Arabidopsis thalia

ALIGNMENTS

RESULT 1

- AAG34077
- ID AAG34077 standard; Protein; 361 AA.
- XX
- AC AAG34077;
- XX
- DT 18-OCT-2000 (first entry)
- XX
- DE Zea mays protein fragment SEQ ID NO: 41410.
- XX
- DE Protein identification; signal transduction pathway; metabolic pathway;
- KW hybridisation assay; genetic mapping; gene expression control; promoter;
- KW termination sequence; corn.
- XX
- OS Zea mays subsp. mays.
- XX
- XX EPI033405-A2.
- XX
- PD 06-SEP-2000.
- XX
- XX 25-FEB-2000; 2000EP-0301439.
- XX
- PR 25-FEB-1999; 99US-0121825.
- PR 05-MAR-1999; 99US-0123180.
- PR 09-MAR-1999; 99US-0123548.
- PR 23-MAR-1999; 99US-0125788.
- PR 25-MAR-1999; 99US-0126264.
- PR 29-MAR-1999; 99US-0126785.
- PR 01-APR-1999; 99US-0127462.
- PR 06-APR-1999; 99US-0128234.
- PR 08-APR-1999; 99US-0128714.
- PR 16-APR-1999; 99US-0129845.
- PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136302.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137242.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140623.  
PR 28-JUN-1999; 99US-0140691.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 18-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149308.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0158299.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.

PR	14-OCT-1999;	99US-0159331.	PN	EP1033405-A2.
PR	14-OCT-1999;	99US-0159637.	XX	
PR	14-OCT-1999;	99US-0159638.	PD	06-SEP-2000.
PR	18-OCT-1999;	99US-0159584.	XX	
PR	21-OCT-1999;	99US-0160741.	PF	
PR	21-OCT-1999;	99US-0160767.	XX	25-FEB-2000; 2000EP-0301439.
PR	21-OCT-1999;	99US-0160768.		
PR	21-OCT-1999;	99US-0160770.	PR	25-FEB-1999; 99US-0121825.
PR	21-OCT-1999;	99US-0160814.	PR	05-MAR-1999; 99US-0123180.
PR	21-OCT-1999;	99US-0160815.	PR	09-MAR-1999; 99US-0123548.
PR	22-OCT-1999;	99US-0160980.	PR	23-MAR-1999; 99US-0125788.
PR	22-OCT-1999;	99US-0160981.	PR	25-MAR-1999; 99US-0126264.
PR	22-OCT-1999;	99US-0160989.	PR	29-MAR-1999; 99US-0126785.
PR	22-OCT-1999;	99US-0161405.	PR	01-APR-1999; 99US-0127462.
PR	25-OCT-1999;	99US-0161404.	PR	06-APR-1999; 99US-0128234.
PR	25-OCT-1999;	99US-0161406.	PR	16-APR-1999; 99US-0128714.
PR	26-OCT-1999;	99US-0161359.	PR	19-APR-1999; 99US-0129845.
PR	26-OCT-1999;	99US-0161360.	PR	19-APR-1999; 99US-0130077.
PR	26-OCT-1999;	99US-0161361.	PR	21-APR-1999; 99US-0130449.
PR	28-OCT-1999;	99US-0161920.	PR	23-APR-1999; 99US-0130510.
PR	28-OCT-1999;	99US-0161992.	PR	23-APR-1999; 99US-0130891.
PR	28-OCT-1999;	99US-0161993.	PR	28-APR-1999; 99US-0131449.
PR	29-OCT-1999;	99US-0162142.	PR	30-APR-1999; 99US-0132048.
Query Match			PR	30-APR-1999; 99US-0132407.
Best Local Similarity 98.4%; Score 1835; DB 21; Length 361;			PR	04-MAY-1999; 99US-0132484.
Matches 355; Conservative 3; Mismatches 3; Indels 0; Gaps 0;			PR	05-MAY-1999; 99US-0132485.
			PR	06-MAY-1999; 99US-0132486.
QY	1	MKALILVGGFGRTRLPRLTSLFPKPLVDFANKPMLHQIEALKEGVTEVVLAINRPEVM 60	PR	06-MAY-1999; 99US-0132487.
Db	1	MKALILVGGFGRTRLPRLTSLFPKPLVDFANKPMLHQIEALKEGVTEVVLAINRPEVM 60	PR	07-MAY-1999; 99US-0132863.
QY	61	INFLKDFEDKLGITITCSQETEPGLTAPLALARDKLAGDGGSPFFVLNSDVISEYPAE 120	PR	11-MAY-1999; 99US-0134256.
Db	61	INFLKDFEDKLGITITCSQETEPGLTAPLALARDKLAGDGGSPFFVLNSDVISEYPAE 120	PR	14-MAY-1999; 99US-0134218.
QY	121	LIFPHKCHGGEATIMTKVDPSKYGVVVMEEATGRVERFEKPKIFVGNKINAGIYLLN 180	PR	14-MAY-1999; 99US-0134221.
Db	121	LIFPHKSHGGEATIMTKVDPSKYGVVVMEEATGRVERFEKPKIFVGNKINAGIYLLN 180	PR	14-MAY-1999; 99US-0134370.
QY	181	PSVLDRELPTSTSEKEVFFQIAADQOLYAMVLPGFWMVQOPRDYITGLRLYLDLSIRKK 240	PR	18-MAY-1999; 99US-0134768.
Db	181	PSVLDRELPTSTSEKEVFFQIAADQOLYAMVLPGFWMVQOPRDYITGLRLYLDLSIRKK 240	PR	19-MAY-1999; 99US-0134941.
QY	241	SAAKLATGARHVGNVLVHESAKIGEGCLIGPDVAIGPCVVVEDGVRLSRTCVMRGVRIKK 300	PR	20-MAY-1999; 99US-0135124.
Db	241	SAAKLAAGARHVGNVLVHESAKIGEGCLIGPDVAIGPCVVVDGVRLSRTCVMRGVRIKK 300	PR	21-MAY-1999; 99US-0135353.
QY	301	HACISNSIIGWHSVVGQWARIENMTILGEDVHVCDVYSNGGVVLPHPHKEIKSSILKPEIV 360	PR	24-MAY-1999; 99US-0135629.
Db	301	HACISNSIIGWHSVVGQWARIENMTILGEDVHVCDVYSNGGVVLPHPHKEIKSSILKPEIV 360	PR	25-MAY-1999; 99US-0136021.
QY	361	M 361	PR	27-MAY-1999; 99US-0136392.
Db	361	M 361	PR	28-MAY-1999; 99US-0136782.
RESULT 2			PR	01-JUN-1999; 99US-0137222.
AAG20009			PR	03-JUN-1999; 99US-0137528.
ID	AAG20009 standard; Protein; 361 AA.		PR	04-JUN-1999; 99US-0137502.
XX			PR	07-JUN-1999; 99US-0137724.
AC			PR	08-JUN-1999; 99US-0138094.
XX			PR	10-JUN-1999; 99US-0138540.
XX			PR	14-JUN-1999; 99US-0138847.
XX			PR	16-JUN-1999; 99US-0139452.
XX			PR	16-JUN-1999; 99US-0139453.
XX			PR	17-JUN-1999; 99US-0139492.
XX			PR	18-JUN-1999; 99US-0139454.
XX			PR	18-JUN-1999; 99US-0139455.
XX			PR	18-JUN-1999; 99US-0139456.
XX			PR	18-JUN-1999; 99US-0139457.
XX			PR	18-JUN-1999; 99US-0139458.
XX			PR	18-JUN-1999; 99US-0139459.
XX			PR	18-JUN-1999; 99US-0139460.
XX			PR	18-JUN-1999; 99US-0139461.
XX			PR	18-JUN-1999; 99US-0139462.
XX			PR	18-JUN-1999; 99US-0139463.
XX			PR	18-JUN-1999; 99US-0139750.
XX			PR	18-JUN-1999; 99US-0139763.
XX			PR	21-JUN-1999; 99US-0139817.
XX			PR	22-JUN-1999; 99US-0139899.
XX			PR	23-JUN-1999; 99US-0140353.
XX			PR	23-JUN-1999; 99US-0140354.
XX			PR	24-JUN-1999; 99US-0140695.
XX			PR	28-JUN-1999; 99US-0140823.
XX			PR	29-JUN-1999; 99US-0140991.
XX			PR	30-JUN-1999; 99US-0141287.
XX			PR	01-JUL-1999; 99US-0141842.

17-OCT-2000 (first entry)

Arabidopsis thaliana protein fragment SEQ ID NO: 22030.

Protein identification; signal transduction pathway; metabolic pathway;  
hybridisation assay; genetic mapping; gene expression control; promoter;  
termination sequence.

Arabidopsis thaliana.

```
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147933.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.

PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 07-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 90.2%; Score 1681; DB 21; Length 361;
Best Local Similarity 88.6%; Pred. No. 1.4e-176;
Matches 320; Conservative 22; Mismatches 19; Indels 0; Gaps 0;

QY 1 MKALILVGEGTRLRPLTSLSPKPLVDFAFKPMILHQLHIEALKEVGYTEVYVLAINRPEYM 60
DD 1 MKALILVGGGTRLRPLTSLSPKPLVDFAFKPMILHQLHIEALKEVGYTEVYVLAINRPEYM 60
QY 61 INFLKDFEDKLGTTTCQSOTTEPLCTAGDIALARDKLDGSCOPFFVLNSDVISSEYFAE 120
DB 61 LNPLKDFETKLEIKITCSOETEPGLTAGPLALARDKLLDGGSGPFFVLNSDVISSEYPLKE 120
QY 121 LKIFHKHCHGEATIMVTKVDEPSKYGVVYMEATGRVEREVEPKIFVGNKINAGIYLLN 180
DB 121 MLEFHKSHGGEASIMVTKVDEPSKYGVVYMEESTGRVEKFEVDPKLYVGNKINAGIYLLN 180
QY 181 PSVLDRIELRPTISIEKEVFPQTAADOOOLYAMVLPGFMMVDVGPDRDYITGLRLYLDSTRKK 240
DB 181 PSVLDRIELRPTISIEKETFPKIAAQQGLYAMVLPGFMMVDIGQPRDYITGLRLYLDSTRKK 240
QY 241 SAAKATGAHVGNVNLVHESAKITGECCLIGDPDVAIGPCGVVEDCVRLSRTVMRGVRIKK 300
DB 241 SPAKLTSGPHVGNVNLVDETTATIGECCLIGDPDVAIGPCGVESGVLRSRTVMRGVRIKK 300
QY 301 HACISNSIIGWSTVSGQWARIENMTILGEDVHVCDVYNSGVVLPKHEKTSSTLKRPEIV 360
DB 301 HACISNSIIGWSTVSGQWARIENMTILGEDVHVSDIYSNGGVVLPKHEKTSNLLKRPEIV 360
QY 361 M 361
DB 361 M 361
```

RESULT 3  
AAG46459  
ID AAG46459 standard; Protein; 361 AA.  
XX  
AC AAG46459;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 58452.  
XX  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PE 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.

```
DR 20-AUG-1999; 99US-0149723.
FR 20-AUG-1999; 99US-0149923.
FR 23-AUG-1999; 99US-0149902.
FR 23-AUG-1999; 99US-0149930.
FR 25-AUG-1999; 99US-0150566.
FR 26-AUG-1999; 99US-0150884.
FR 27-AUG-1999; 99US-0151065.
FR 27-AUG-1999; 99US-0151066.
FR 27-AUG-1999; 99US-0151080.
FR 30-AUG-1999; 99US-0151303.
FR 31-AUG-1999; 99US-0151438.
FR 01-SEP-1999; 99US-0151930.
FR 07-SEP-1999; 99US-0152363.
FR 10-SEP-1999; 99US-0153070.
FR 13-SEP-1999; 99US-0153758.
FR 15-SEP-1999; 99US-0154018.
FR 16-SEP-1999; 99US-0154039.
FR 20-SEP-1999; 99US-0154779.
FR 22-SEP-1999; 99US-0155139.
FR 23-SEP-1999; 99US-0155486.
FR 24-SEP-1999; 99US-0155659.
FR 28-SEP-1999; 99US-0156458.
FR 29-SEP-1999; 99US-0156596.
FR 04-OCT-1999; 99US-0157117.
FR 05-OCT-1999; 99US-0157753.
FR 06-OCT-1999; 99US-0157865.
FR 07-OCT-1999; 99US-0158029.
FR 08-OCT-1999; 99US-0158212.
FR 12-OCT-1999; 99US-0158369.
FR 13-OCT-1999; 99US-0159293.
FR 13-OCT-1999; 99US-0159294.
FR 13-OCT-1999; 99US-0159295.
FR 14-OCT-1999; 99US-0159329.
FR 14-OCT-1999; 99US-0159330.
DR 14-OCT-1999; 99US-0159331.
FR 14-OCT-1999; 99US-0159637.
FR 14-OCT-1999; 99US-0159638.
DR 18-OCT-1999; 99US-0159584.
FR 21-OCT-1999; 99US-0160741.
FR 21-OCT-1999; 99US-0160767.
DR 21-OCT-1999; 99US-0160768.
FR 21-OCT-1999; 99US-0160770.
FR 21-OCT-1999; 99US-0160814.
FR 21-OCT-1999; 99US-0160815.
FR 22-OCT-1999; 99US-0160980.
FR 22-OCT-1999; 99US-0160981.
DR 22-OCT-1999; 99US-0160989.
FR 25-OCT-1999; 99US-0161404.
FR 25-OCT-1999; 99US-0161405.
FR 25-OCT-1999; 99US-0161406.
FR 26-OCT-1999; 99US-0161359.
FR 26-OCT-1999; 99US-0161360.
FR 26-OCT-1999; 99US-0161361.
FR 28-OCT-1999; 99US-0161920.
FR 28-OCT-1999; 99US-0161992.
FR 28-OCT-1999; 99US-0161993.
FR 29-OCT-1999; 99US-0162142.

Query Match 90.2%; Score 1681; DB 21; Length 361;
Best Local Similarity 88.6%; Pred. No. 1.4e-176;
Matches 320; Conservative 22; Mismatches 19; Indels 0; Gaps 0;

Qy 1 MKALILVGGGCTRLRLPLTSPFKPLVDFAFKPMILHQIALKEVGVTEVLAINRPEVM 60
Db 1 MKALILVGGGCTRLRLPLTSPFKPLVDFAFKPMILHQIALKEVGVTEVLAINRPEVM 60

Qy 61 INFLKDFEDKLGITIGTSQETPERLGTAGLALARDKLDGSGQPFVFLNSDVISEYPFAE 120
Db 61 LNFLKDFETKLEIKTSQETPEPLGTAGLALARDKLDGSGEFVFLNSDVISEYPLKE 120

Qy 121 LIKPHKHGGGATIMVTKVDPSYGVVVMEEATGRVERFVEKPKIFVGNKINAGIYLLN 180
Db 121 MLEPHKHGGGASIMVTKVDPSYGVVVMEEATGRVEKPKLYVGNKINAGIYLLN 180
```

```
QY 181 PSVLDRIELRPTSIKREVFQIAADQOLXAMYLPGFWMVDYQGPDXITIGURLYDLSLRKK 240
Db 181 PSVLDRIELRPTSIKRETFEKIAAAGLYAMVLPGFMDIGQPRDVTITGLRLYDLSLRKK 240
QY 241 SAAKLATGAHVGNVNLVHESAKIGEGCLIGDPVAIGPGCVVDEGVRLSRCTVMRGVRIKK 300
Db 241 SPAKLTSGPHIVGNVLVDETATIGEGCLIGDPVAIGPGCIVESGVRLSRCTVMRGVRIKK 300
QY 301 HACISNSIICWHSTVQOWARIENKMTILGEDVHVCDVSNVGGVVLPHKETKSSILKPEIV 360
Db 301 HACISSIICWHSTVQOWARIENMTILGEDVHVSDVETYSNGGVVLPHKETKSNILKPEIV 360
QY 361 M 361
Db 361 M 361

RESULT 4
AAV77936
ID AAV77936 standard; protein; 361 AA.
XX
AC AAV77936;
XX
DT 14-JUN-2000 (first entry)
XX
DE A. thaliana environmental stress tolerance related protein.
KW Environmental stress; plant; transgenic plant; anaerobic; flooding; cold;
  dehydration; drought; heat stress; salinity; osmotolerance.
OS Arabidopsis thaliana.
XX
DN WO2000008187-A2.
XX
PD 17-FEB-2000.
XX
PF 04-AUG-1999; 99WO-EP05652.
XX
PR 04-AUG-1999; 98EP-0202634.
XX
PA (VLAAs) YLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
PI Lee JH, Verbruggen N;
XX
XX WPI: 2000-205726/18.
DR N-PEDB; AAZ98316.
XX
XX Isolation of polynucleic acids useful for producing transgenic plant by
  isolating genes involved in tolerance to environmental stress -
  Claim 12; Page 127-129; 312pp; English.
CC The invention relates to isolation of coding sequences and/or genes
  involved in tolerance to environmental stress in plants. The sequences
  (AAZ98305-298365) are useful for producing a transgenic plant having
  enhanced tolerance or resistance to environmental stress conditions such
  as anaerobic, flooding, cold, dehydration, drought, heat stress or
  salinity. This is useful for producing improved yield, growth,
  development and productivity under environmental stress conditions, and
  also provides growth of crops in areas where they cannot grow without
  the induced osmotolerance. Sequences AAV7925-984 represent polypeptide
  sequences from A. thaliana that are encoded by the genes involved in
  environmental stress tolerance.
XX
SQ Sequence 361 AA;
```

```
Query Match 90.2%; Score 1681; DB 21; Length 361;
Best Local Similarity 88.6%; Pred. No. 1.4e-176;
Matches 320; Conservative 22; Mismatches 19; Indels 0; Gaps 0;

Qy 1 MKALILVGGGCTRLRLPLTSPFKPLVDFAFKPMILHQIALKEVGVTEVLAINRPEVM 60
Db 1 MKALILVGGGCTRLRLPLTSPFKPLVDFAFKPMILHQIALKEVGVTEVLAINRPEVM 60
```





PR	30-APR-1999;	99US-0132407.	PR	21-JUL-1999;	99US-0145088.
PR	04-MAY-1999;	99US-0132484.	PR	22-JUL-1999;	99US-0145085.
PR	05-MAY-1999;	99US-0132485.	PR	22-JUL-1999;	99US-0145087.
PR	06-MAY-1999;	99US-0132486.	PR	22-JUL-1999;	99US-0145089.
PR	07-MAY-1999;	99US-0132487.	PR	22-JUL-1999;	99US-0145192.
PR	07-MAY-1999;	99US-0132863.	PR	23-JUL-1999;	99US-0145145.
PR	11-MAY-1999;	99US-0134256.	PR	23-JUL-1999;	99US-0145218.
PR	14-MAY-1999;	99US-0134218.	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	99US-0134219.	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	99US-0134221.	PR	27-JUL-1000;	99US-0145913.
PR	14-MAY-1999;	99US-0134370.	PR	27-JUL-1999;	99US-0145918.
PR	18-MAY-1999;	99US-0134768.	PR	27-JUL-1999;	99US-0145919.
PR	19-MAY-1999;	99US-0134941.	PR	28-JUL-1999;	99US-0145951.
PR	20-MAY-1999;	99US-0135124.	PR	02-AUG-1999;	99US-0146386.
PR	21-MAY-1999;	99US-0135353.	PR	02-AUG-1999;	99US-0146388.
PR	24-MAY-1999;	99US-0135629.	PR	02-AUG-1999;	99US-0146389.
PR	25-MAY-1999;	99US-0136021.	PR	03-AUG-1999;	99US-0147038.
PR	27-MAY-1999;	99US-0136392.	PR	04-AUG-1999;	99US-0147204.
PR	28-MAY-1999;	99US-0136782.	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	99US-0137222.	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	99US-0137528.	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	99US-0137502.	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	99US-0137724.	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	99US-0138094.	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	99US-0138540.	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	99US-0138847.	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	99US-0139119.	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	99US-0139452.	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	99US-0139453.	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	99US-0139492.	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	99US-0139454.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139462.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	99US-0139899.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0142154.	PR	20-SEP-1999;	99US-0154779.
PR	02-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155139.
PR	06-JUL-1999;	99US-0142390.	PR	23-SEP-1999;	99US-0155486.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-0155659.
PR	09-JUL-1999;	99US-0142920.	PR	28-SEP-1000;	99US-0156458.
PR	12-JUL-1999;	99US-0142977.	PR	29-SEP-1999;	99US-0156596.
PR	13-JUL-1999;	99US-0143542.	PR	04-OCT-1999;	99US-0157117.
PR	14-JUL-1999;	99US-0143624.	PR	05-OCT-1999;	99US-0157753.
PR	15-JUL-1999;	99US-0144005.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144085.	PR	07-OCT-1999;	99US-0158029.
PR	16-JUL-1999;	99US-0144086.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144325.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159329.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159637.
PR	20-JUL-1999;	99US-0144884.	PR	14-OCT-1999;	99US-0159638.
PR	21-JUL-1999;	99US-0144814.	PR	18-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160741.

[illegible]

PR 16-JUL-1999;	99US-0144085.	PR 07-OCT-1999;	99US-0158029.
PR 16-JUL-1999;	99US-0144086.	PR 08-OCT-1999;	99US-0158232.
PR 19-JUL-1999;	99US-0144325.	PR 12-OCT-1999;	99US-0158369.
PR 10-JUL-1999;	99US-0144331.	PR 13-OCT-1999;	99US-0159209.
PR 19-JUL-1999;	99US-0144332.	PR 13-OCT-1999;	99US-0159294.
PR 19-JUL-1999;	99US-0144333.	PR 13-OCT-1999;	99US-0159295.
PR 19-JUL-1999;	99US-0144334.	PR 14-OCT-1999;	99US-0159329.
PR 19-JUL-1999;	99US-0144335.	PR 14-OCT-1999;	99US-0159330.
PR 20-JUL-1999;	99US-0144352.	PR 14-OCT-1999;	99US-0159331.
PR 20-JUL-1999;	99US-0144632.	PR 14-OCT-1999;	99US-0159637.
PR 20-JUL-1999;	99US-0144804.	PR 14-OCT-1999;	99US-0159638.
PR 21-JUL-1999;	99US-0144814.	PR 18-OCT-1999;	99US-0159584.
PR 21-JUL-1999;	99US-0145086.	PR 21-OCT-1999;	99US-0160741.
PR 21-JUL-1999;	99US-0145088.	PR 21-OCT-1999;	99US-0160767.
PR 22-JUL-1999;	99US-0145085.	PR 21-OCT-1999;	99US-0160768.
PR 22-JUL-1999;	99US-0145087.	PR 21-OCT-1999;	99US-0160770.
PR 22-JUL-1999;	99US-0145089.	PR 21-OCT-1999;	99US-0160814.
PR 22-JUL-1999;	99US-0145192.	PR 21-OCT-1999;	99US-0160815.
PR 23-JUL-1999;	99US-0145145.	PR 22-OCT-1999;	99US-0160980.
PR 23-JUL-1999;	99US-0145218.	PR 22-OCT-1999;	99US-0160981.
PR 23-JUL-1999;	99US-0145224.	PR 22-OCT-1999;	99US-0160989.
PR 26-JUL-1999;	99US-0145276.	PR 25-OCT-1999;	99US-0161404.
PR 27-JUL-1999;	99US-0145913.	PR 25-OCT-1999;	99US-0161405.
PR 27-JUL-1999;	99US-0145918.	PR 25-OCT-1999;	99US-0161406.
PR 27-JUL-1999;	99US-0145919.	PR 26-OCT-1999;	99US-0161359.
PR 28-JUL-1999;	99US-0145931.	PR 26-OCT-1999;	99US-0161360.
PR 02-AUG-1999;	99US-0146386.	PR 26-OCT-1999;	99US-0161361.
PR 02-AUG-1999;	99US-0146388.	PR 28-OCT-1999;	99US-0161920.
PR 02-AUG-1999;	99US-0146389.	PR 28-OCT-1999;	99US-0161992.
PR 03-AUG-1999;	99US-0147038.	PR 28-OCT-1999;	99US-0161993.
PR 04-AUG-1999;	99US-0147204.	PR 29-OCT-1999;	99US-0162142.
PR 04-AUG-1999;	99US-0147302.		
PR 05-AUG-1999;	99US-0147192.		
PR 06-AUG-1999;	99US-0147260.		
PR 06-AUG-1999;	99US-0147303.		
PR 06-AUG-1999;	99US-0147416.		
PR 09-AUG-1999;	99US-0147493.		
PR 09-AUG-1999;	99US-0147935.		
PR 10-AUG-1999;	99US-0148171.		
PR 12-AUG-1999;	99US-0148319.		
PR 13-AUG-1999;	99US-0148341.		
PR 13-AUG-1999;	99US-0148565.		
PR 13-AUG-1999;	99US-0148684.		
PR 16-AUG-1999;	99US-0149368.		
PR 17-AUG-1999;	99US-0149175.		
PR 18-AUG-1999;	99US-0149426.		
PR 20-AUG-1999;	99US-0149722.		
PR 20-AUG-1999;	99US-0149723.		
PR 20-AUG-1999;	99US-0149929.		
PR 23-AUG-1999;	99US-0149902.		
PR 23-AUG-1999;	99US-0149930.		
PR 23-AUG-1999;	99US-0150566.		
PR 26-AUG-1999;	99US-0150884.		
PR 27-AUG-1999;	99US-0151065.		
PR 27-AUG-1999;	99US-0151066.		
PR 27-AUG-1999;	99US-0151080.		
PR 30-AUG-1999;	99US-0151303.		
PR 31-AUG-1999;	99US-0151438.		
PR 01-SEP-1999;	99US-0151930.		
PR 07-SEP-1999;	99US-0152363.		
PR 10-SEP-1999;	99US-0153070.		
PR 13-SEP-1999;	99US-0153758.		
PR 13-SEP-1999;	99US-0154018.		
PR 16-SEP-1999;	99US-0154039.		
PR 20-SEP-1999;	99US-0154779.		
PR 22-SEP-1999;	99US-0155139.		
PR 23-SEP-1999;	99US-0155139.		
PR 23-SEP-1999;	99US-0155486.		
PR 24-SEP-1999;	99US-0155659.		
PR 28-SEP-1999;	99US-0156458.		
PR 29-SEP-1999;	99US-0156596.		
PR 04-OCT-1999;	99US-0157117.		
PR 05-OCT-1999;	99US-0157753.		
PR 06-OCT-1999;	99US-0157865.		
<hr/>			
Query Match 82.6%; Score 1540; DB 21; Length 302;			
Best Local Similarity 98.0%; Pred. No. 3.9e-161;			
Matches 296; Conservative 3; Mismatches 3; Indels 0; Gaps 0;			
Qy 60	MINFLKDFEDKLGITITCSQETPLCTAGLALARKLADGSGQPFVLNSDVISEYPPFA 119		
Db 1	MINFLKDFEDKLGITITCSQETPLCTAGLALARKLADGSGEPFVLNSDVISEYPPFA 60		
Qy 120	ELIKFKHCHGGEATIMVTKVDEPSKYGVVYVMEBEATGRVERFEKPIFVGNKINAGTYLL 179		
Db 61	ELIKFKHCHGGEATIMVTKVDEPSKYGVVYVMEBEATGRVERFEKPIFVGNKINAGTYLL 120		
Qy 180	NPSVLDRIELRETSIEKEVFPQIADQQLYAMVLPFGVNDVGPQDYITGLRLYLSIRK 239		
Db 121	NPSVLDRIELRETSIEKEVFPQIADQQLYAMVLPFGVNDVGPQDYITGLRLYLSIRK 180		
Qy 240	KSAAKLATGAHVGVNVLVHESAKIGRGCLIGPDVAIGPCVWEDGVRLSCTVMRGVRIK 299		
Db 181	KSAARLAAGAHVGVNVLVHESAKIGRGCLIGPDVAIGPCVYADGVRLSCTVMRGVRIK 240		
Qy 300	KHACISNSTIGHSTVGOWARIENMTILGEDVHVCDEVYNSGGVVLPHKEIKSSILKPEI 359		
Db 241	KHACISNSTIGHSTVGOWARIENMTILGEDVHVCDEVYNSGGVVLPHKEIKSSILKPEI 300		
Qy 360	VM 361		
Db 301	VM 302		
RESULT 0			
AAG20010			
ID	AAG20010 standard; Protein; 329 AA.		
XX			
AC	AAG20010;		
XX			
DT	17-OCT-2000 (first entry)		
XX			
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 22031.		
XX			
KW	protein identification; signal transduction pathway; metabolic pathway;		
KW	hybridisation assay; genetic mapping; gene expression control; promoter;		

KW	termination sequence.	PR	28-JUN-1999;	99US-0140823.
XX		PR	29-JUN-1999;	99US-0140991.
OS	Arabidopsis thaliana.	PR	30-JUN-1999;	99US-0141287.
XX		PR	01-JUL-1999;	99US-0141842.
PN	EPI033405-A2.	PR	01-JUL-1999;	99US-0142154.
XX		PR	02-JUL-1999;	99US-0142055.
PD	06-SEP-2000.	PR	06-JUL-1999;	99US-0142390.
XX		PR	08-JUL-1999;	99US-0142803.
PF	25-FEB-2000; 2000EP-0301439.	PR	09-JUL-1999;	99US-0142920.
XX		PR	12-JUL-1999;	99US-0142977.
PR	25-FEB-1999;	PR	13-JUL-1999;	99US-0143542.
PR	05-MAR-1999;	PR	14-JUL-1999;	99US-0143624.
PR	09-MAR-1999;	PR	15-JUL-1999;	99US-0144005.
PR	23-MAR-1999;	PR	16-JUL-1999;	99US-0144085.
PR	25-MAR-1999;	PR	16-JUL-1999;	99US-0144086.
PR	29-MAR-1999;	PR	19-JUL-1999;	99US-0144325.
PR	01-APR-1999;	PR	19-JUL-1999;	99US-0144331.
PR	06-APR-1999;	PR	19-JUL-1999;	99US-0144332.
PR	08-APR-1999;	PR	19-JUL-1999;	99US-0144333.
PR	16-APR-1999;	PR	19-JUL-1999;	99US-0144334.
PR	19-APR-1999;	PR	19-JUL-1999;	99US-0144335.
PR	21-APR-1999;	PR	20-JUL-1999;	99US-0144352.
PR	23-APR-1999;	PR	20-JUL-1999;	99US-0144632.
PR	23-APR-1999;	PR	20-JUL-1999;	99US-0144884.
PR	28-APR-1999;	PR	21-JUL-1999;	99US-0144814.
PR	30-APR-1999;	PR	21-JUL-1999;	99US-0145086.
PR	30-APR-1999;	PR	21-JUL-1999;	99US-0145088.
PR	04-MAY-1999;	PR	22-JUL-1999;	99US-0145085.
PR	05-MAY-1999;	PR	22-JUL-1999;	99US-0145087.
PR	06-MAY-1999;	PR	22-JUL-1999;	99US-0145089.
PR	08-MAY-1999;	PR	22-JUL-1999;	99US-0145192.
PR	07-MAY-1999;	PR	23-JUL-1999;	99US-0145145.
PR	11-MAY-1999;	PR	23-JUL-1999;	99US-0145218.
PR	14-MAY-1999;	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145918.
PR	18-MAY-1999;	PR	27-JUL-1999;	99US-0145919.
PR	19-MAY-1999;	PR	28-JUL-1999;	99US-0145951.
PR	20-MAY-1999;	PR	02-AUG-1999;	99US-0146386.
PR	21-MAY-1999;	PR	02-AUG-1999;	99US-0146388.
PR	24-MAY-1999;	PR	02-AUG-1999;	99US-0146389.
PR	25-MAY-1999;	PR	03-AUG-1999;	99US-0147038.
PR	27-MAY-1999;	PR	04-AUG-1999;	99US-0147204.
PR	28-MAY-1999;	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	PR	31-AUG-1999;	99US-01513438.
PR	23-JUN-1999;	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	PR	07-SEP-1999;	99US-0152363.











PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140333.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140931.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142330.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145152.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149920.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 31-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 14-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.



RESULT 13  
AAV79210  
ID AAV79210 standard; Protein; 360 AA.  
XX  
AC AAV79210;  
XX  
DT 19-JUN-2000 (first entry)  
XX  
DE Human transferase TRNSFS-2.  
XX  
KW Transferase; TRNSFS-2; human; antitumour; cancer;  
KW gastrointestinal disorder; developmental disorder;  
KW genetic disorder; neurological disorder; reproductive disorder;  
KW smooth muscle disorder; immunological disorder; inflammation;  
KW diagnosis; therapy; mannose-1-phosphate guanyltransferase.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 290  
FT Modified-site /note= "potential O-phosphorylation"  
FT Modified-site /note= "potential O-phosphorylation"  
FT Modified-site 136  
FT Modified-site /note= "potential O-phosphorylation"  
FT Modified-site 191  
FT Modified-site /note= "potential O-phosphorylation"  
FT Modified-site 301  
FT Modified-site /note= "potential O-phosphorylation"  
FT Modified-site 348  
FT Modified-site /note= "potential O-phosphorylation"  
FT Modified-site 352  
FT Modified-site /note= "potential O-phosphorylation"  
FT Modified-site 21  
FT Modified-site /note= "potential O-phosphorylation"  
FT Modified-site 153  
FT Modified-site /note= "potential O-phosphorylation"  
FT Modified-site 168  
FT Modified-site /note= "potential O-phosphorylation"  
FT Modified-site 235  
FT Modified-site /note= "potential O-phosphorylation"  
FT Modified-site 259  
FT Modified-site /note= "potential O-phosphorylation"  
FT Modified-site 144  
FT Modified-site /note= "potential O-phosphorylation"  
FT Modified-site 265  
FT Modified-site /note= "potential N-glycosylation"  
FT Modified-site 271  
FT Modified-site /note= "potential N-glycosylation"  
FT Modified-site 322  
FT Modified-site /note= "potential N-glycosylation"  
FT Peptide 256..284  
FT Peptide /note= "bacterial hexapeptide transferase signature"  
FT Peptide 7..44  
FT Peptide /note= "putative ADP-glucose  
FT Peptide 106..144  
FT Peptide /note= "putative ADP-glucose  
FT Peptide /note= "pyrophosphorylase signature"  
XX  
PN WO200014251-A2.  
XX  
PD 16-MAR-2000.  
XX  
PE 09-SEP-1999; 99WO-US20989.  
XX  
PR 10-SEP-1998; 98US-0150657.  
PR 04-NOV-1998; 98US-0186779.  
PR 11-MAY-1999; 99us-0133642.  
XX

PA (INCY-) INCYTE PHARM INC.  
XX Tang YT, Corley NC, Guegler KJ, Baughn MR, Lal P, Yue H;  
PI Hillman JL, Azimzai Y;  
XX  
DR WPI: 2000-256996/22.  
DR N-PSDB; AAZ94202.  
XX  
PT Human transferase proteins useful for preventing, diagnosing and  
PT treating cancers and developmental, gastrointestinal, genetic,  
PT immunological, neurological, reproductive and smooth muscle disorders -  
XX  
PS Claim 1; Page 80; 113pp; English.  
XX  
CC The present sequence is that of human transferase TRNSFS-2, 1 of  
CC 15 claimed novel human transferase proteins of the invention (see  
CC AAY79209-23). The sequence was deduced from a cDNA clone (see  
CC AAY79202) isolated from an ileum tissue library. It shows homology  
CC to mannose-1-phosphate guanyltransferases. TRNSFS-2 is expressed  
CC in reproductive and gastrointestinal tissues, especially those  
CC associated with cancer and inflammation. The new human transferase  
CC proteins and the polynucleotides encoding them can be used in the  
CC diagnosis, prevention and treatment of cancer, developmental  
CC disorders, gastrointestinal disorders, genetic disorders,  
CC immunological disorders, neurological disorders, reproductive  
CC disorders, and smooth muscle disorders. The polypeptides can also  
CC be used to raise antibodies, and to screen for agonists and  
CC antagonists of transferase activity.  
XX  
SO Sequence 360 AA;  
Query Match 63.5%; Score 1183.5; DB 21; Length 360;  
Best Local Similarity 61.2%; Pred. No. 1.3e-121;  
Matches 221; Conservative 65; Mismatches 74; Indels 1; Gaps 1:  
QY 1 MKALILVGGFGTRLRPLTLSTFPKPLVDFAANKPMLHIOIEALKEVGTVVLAINYRPEVM 60  
DB 1 MKALILVGGYGTRLRPLTLSTPKPLVDFAANKPMLHIOIEALKEVGTVVLAINYRPEVM 60  
QY 61 INFLKDPEDKIGITTCSTOFTPLTAGPLALARDKADCGOPFFVLNSDVISEYDFAE 120  
DB 61 EKEMKAOEORLIGIRISMSHEEPLGTAGPLALARDLISE-TADFFVLNSDVICDFPQA 119  
QY 121 LKFKHKGHEATIMVTKYDPSKYGVVYMEATGRVREVERKPFVGNKINAGIYLLN 180  
DB 120 MVQFHRHGGESILVTKVEEPSKYGVVYCEADTGRHREVERKPFVGNKINAGIYLLN 179  
QY 181 PSVLDRIELRPTSTIEKEVFPQIAADQOLYAMVLPGFWMQVQPRDYITGLRLYLDSTRKK 240  
DB 180 PAVLRRIQLPTSTIEKEVFPIMAKEGLYAMELQGFWMQVQPRDYITGLRLYLDSTRKK 239  
QY 241 SAAKLATGAHVGNVNLVHESAKIGEGCLIGPDVAIGPCVVEDGVRSLRCTVMRGVRIKK 300  
DB 240 QPERLCSPGIVGNVLVDFPSARIGQNCISGPNVSLGPGVYVEDGVCIRRTCTVLRDAIRS 299  
QY 301 HACISNIIIGWSTVGQWARIENNTILGDEHVHCVDSYNGVVLPHKETSSTLKPEIV 360  
DB 300 HSWLESCIVGWRRCRGQWVRMENVTVLGEDVIYNDELYLNGASVLPKHSIGESVPEPRII 359  
QY 361 M 361  
DB 360 M 360  
RESULT 14  
AAV52357  
ID AAV52357 standard; Protein; 363 AA.  
XX  
AC AAV52357;  
XX  
DT 25-JAN-2002 (first entry)  
XX  
DE NDP-hexose pyrophosphorylase homolog.

XX Geminivirus; plant; viral infection; transgenic plant;  
KW tomato yellow leaf curl virus.  
XX Schizosaccharomyces pombe.  
XX FR2806095-A1.  
XX 14-SEP-2001.  
XX 10-MAR-2000; 2000FR-00031140.  
XX 10-MAR-2000; 2000FR-00031140.  
XX (GENT-) GENTECH SARL.  
XX Bejarano ER, Castillo GA, Collinet D, Donoso CI, Iniesta JR;  
PI Grevesse C, Hericourt F;  
XX WPI; 2001-628275/73.  
DR N-PSDB; ABA01227.  
XX New polynucleotides for producing transgenic plants resistant to  
PT geminivirus infection comprising polynucleotides encoding proteins  
PT which interact with at least one of the products of the geminivirus  
PT genome -  
XX Claim 4; Pages 38-39; 106pp; French.  
XX The present invention relates to coding sequences encoding proteins which  
CC interact with at least one of the six products of the geminivirus genome  
CC necessary for infection of a plant by the virus. The present sequence is  
CC one such protein. The coding sequences are useful for producing  
CC transgenic plants resistant to geminivirus infection, particularly tomato  
CC yellow leaf curl virus.  
XX Sequence 363 AA;  
Query Match 63.0%; Score 1175; DB 22; Length 363;  
Best Local Similarity 62.1%; Pred. No. 1.1e-120;  
Matches 226; Conservative 54; Mismatches 80; Indels 4; Gaps 4;  
QY 1 MKALILVGGFGRTRPLTLSPKPLVDFAFKPMILHQLKEVGTWVWVLAIVRPEVM 60  
DB 1 MKALILVGGFGRTRPLTLSPKPLVDFAFKPMILHQLKEVGTWVWVLAIVRPEIM 60  
QY 61 INFLKDFEDKLGITITCSQETEPGLTAGPLALARDKLDAGSGQPFVFLNSDVISEYPAE 120  
DB 61 VEALKYKEYNVNITFSENEPLGTAGPLALARDILAK-DHSPFFVFLNSDVISEYPAE 119  
QY 121 LIKFKHCHGEATIMVTKVDEPSKYGVV-MEEATGRVERFVEKPIFVGNKINAGIYLL 179  
DB 120 LAAPHKAGAGGTIVTKVEEPSKYGVVHVYVNSLESIERFVEKPIFVGNKINAGIYIL 179  
QY 180 NPSVLDRIELRPTSTIEKEVFQIAADQOLYAMVLPFGWMDYQPRDYITGLRLYLDLSIRK 239  
DB 180 NPSVLDRIELRPTSTIEKEVFQIAADQOLYAMVLPFGWMDYQPRDYITGLRLYLDLSIRK 239  
QY 240 KSAAKLA-TGAHVGNVNLVHESAKIGEGCLIGDPDAIGPGCVVEDGVRLSRCTVMRGVRI 298  
DB 240 HKPEILAPASSNIIGNVLIDPSATIGNCKIGPNVIGPNVTIGDGVRLQRCAILKSSRV 299  
QY 299 KKHACINSIIGWISTVQWARIENMVLIGDGVHVCDEYVNGGVLPHEIKSSILKP- 357  
DB 300 RDHAWKSSIVGNWSTLGSRLNVSVLGDDVVVNDIYVNGGSLPHKHSISANIEVPG 359  
QY 358 EIVM 361  
DB 360 TIVM 363  
RESULT 15  
ABB57869

ID ABB57869 standard; Protein; 369 AA.  
XX  
AC ABB57869;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 399.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US09231.  
XX  
PR 23-MAR-2000; 2000US-191637P.  
PR 11-JUL-2000; 2000US-0614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX  
DR WPI; 2001-656860/75.  
DR N-PSDB; ABL01972.  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
XX  
PS Disclosure; SEQ ID NO 399; 21pp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins  
CC (ABB57737-ABB72072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pot\_sequences.  
XX  
SQ Sequence 369 AA;  
Query Match 62.9%; Score 1173; DB 22; Length 369;  
Best Local Similarity 61.7%; Pred. No. 1.9e-120;  
Matches 222; Conservative 64; Mismatches 72; Indels 2; Gaps 2;  
QY 2 KALILVGGFGRTRPLTLSPKPLVDFAFKPMILHQLKEVGTWVWVLAIVRPEVMI 61  
DB 12 RALILVGGFGRTRPLTLSPKPLVDFAFKPMILHQLKEVGTWVWVLAIVRPEVMI 71  
QY 62 NFKLDFEDKLGITITCSQETEPGLTAGPLALARDKLDAGSGQPFVFLNSDVISEYPAEL 121  
DB 72 KELKVEAKLGLVELIFSHETEPGLTAGPLALAKTILA-ASSEFFVFLNSDVICDFPKQL 130  
QY 122 IFKFKHCHGEATIMVTKVDEPSKYGVVMEATGRVERFVEKPIFVGNKINAGIYLLNP 181  
DB 131 VQFHCNHGEGTIVTKVEEPSKYGVVLYDE-NGCINKFTEKPEQEFVSNKINAGIYIFNP 189  
QY 182 SVLDRIELRPTSTIEKEVFQIAADQOLYAMVLPFGWMDYQPRDYITGLRLYLDLSIRKS 241  
DB 190 SVLDRIELRPTSTIEKEVFQIAADQOLYAMVLPFGWMDYQPRDYITGLRLYLDLSIRKS 249  
QY 242 AAKLATGAHVGNVNLVHESAKIGEGCLIGDPDAIGPGCVVEDGVRLSRCTVMRGVRIKKH 301  
DB 250 SPKLYTGPVGNVNLVDPTAKIGEGCGRIGPNVTIGDGVIEDGVCIKRSILKAGIYVRSH 309

Search completed: November 27, 2002, 05:07:04  
Job time : 70 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 05:07:12 ; Search time 26 Seconds  
(without alignments)  
408.526 Million cell updates/sec

Title: US-09-374-967-2  
Perfect score: 1864  
Sequence: 1 MKALLVGGFTRLRLPLTSL.....GVVLPHEIKSSILKPEIVM 361

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A-COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B-COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A-COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B-COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS-COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	% Match	Query Length	DB ID	Description
1	1178	63.2	361	3	US-09-032-372-12 Sequence 12, Appl
2	401	21.5	421	3	US-09-032-372-3 Sequence 3, Appli
3	307	16.5	355	1	US-08-196-218-34 Sequence 34, Appl
4	307	16.5	355	1	US-08-681-953-34 Sequence 34, Appl
5	296	15.9	355	4	US-09-194-905-10 Sequence 10, Appl
6	241.5	13.0	461	4	US-09-134-001C-3316 Sequence 3316, Ap
7	239	12.8	146	4	US-09-370-838-190 Sequence 190, App
8	228.5	12.3	224	4	US-09-199-637A-175 Sequence 175, App
9	193	10.4	431	1	US-08-469-202-14 Sequence 14, Appl
10	193	10.4	431	1	US-08-484-434C-14 Sequence 14, Appl
11	191	10.2	431	1	US-08-090-523-4 Sequence 4, Appli
12	191	10.2	431	1	US-08-398-627-4 Sequence 4, Appli
13	191	10.2	431	1	US-08-406-858-4 Sequence 4, Appli
14	191	10.2	431	5	PCT-US91-04036-4 Sequence 4, Appli
15	191	10.2	431	5	PCT-US91-05275-4 Sequence 4, Appli
16	191	10.2	488	4	US-09-444-728-2 Sequence 2, Appli
17	190	10.2	292	3	US-09-320-878-14 Sequence 14, Appl
18	190	10.2	292	4	US-09-105-537-12 Sequence 12, Appl
19	190	10.2	3782	4	US-09-105-537-4 Sequence 4, Appli
20	188	10.1	431	1	US-08-090-523-2 Sequence 2, Appli
21	188	10.1	431	1	US-08-398-627-2 Sequence 2, Appli
22	188	10.1	431	1	US-08-406-858-2 Sequence 2, Appli
23	188	10.1	431	5	PCT-US91-04036-2 Sequence 2, Appli
24	188	10.1	431	5	PCT-US91-05275-2 Sequence 2, Appli
25	186.5	10.0	518	1	US-08-485-241-3 Sequence 3, Appli
26	186.5	10.0	518	2	US-08-874-162-3 Sequence 3, Appli
27	186	10.0	517	1	US-08-485-241-5 Sequence 5, Appli

## SUMMARIES

RESULT 1  
US-09-032-372-12  
; Sequence 12, Application US/09032372  
; Patent No. 6008337  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Yue, Henry  
; APPLICANT: Lal, Preeti  
; TITLE OF INVENTION: CELL CYCLE RELATED PROTEINS  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/032,372  
; FILING DATE: Herewith  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0478 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-855-0555  
; TELEFAX: 650-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 361 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GenBank  
; CLONE: 1292898  
US-09-032-372-12

## ALIGNMENTS

28	186	10.0	517	2	US-08-874-162-5	Sequence 5, Appli
29	186	10.0	518	1	US-08-299-675-2	Sequence 2, Appli
30	179.5	9.6	229	3	US-09-024-023-2	Sequence 2, Appli
31	179.5	9.6	229	4	US-09-531-111-2	Sequence 2, Appli
32	174	9.3	459	3	US-08-971-782-2	Sequence 2, Appli
33	174	9.3	459	4	US-09-309-026-2	Sequence 2, Appli
34	169.5	9.1	471	2	US-08-535-276-4	Sequence 4, Appli
35	169.5	9.1	471	4	US-09-335-234-4	Sequence 4, Appli
36	168.5	9.0	295	4	US-09-134-001C-4466	Sequence 4466, Ap
37	166.5	8.9	507	1	US-08-097-829-6	Sequence 6, Appli
38	166.5	8.9	507	1	US-08-577-403-6	Sequence 6, Appli
39	163	8.7	95	3	US-09-024-023-4	Sequence 4, Appli
40	163	8.7	95	4	US-09-531-111-4	Sequence 4, Appli
41	163	8.7	521	1	US-08-090-523-8	Sequence 8, Appli
42	163	8.7	521	1	US-08-398-627-8	Sequence 8, Appli
43	163	8.7	521	1	US-08-406-858-8	Sequence 8, Appli
44	163	8.7	521	5	PCT-US91-04036-8	Sequence 8, Appli
45	163	8.7	521	5	PCT-US94-05275-8	Sequence 8, Appli

Query Match 63.2%; Score 1178; DB 3; Length 361;  
Best Local Similarity 61.0%; Pred. NO. 6.5e-123;  
Matches 221; Conservative 56; Mismatches 83; Indels 2; Gaps 2;  
QY 1 MKAILVGGGTRLRPLTLSPKPLVDFANKPMLHIOEALKEVGVTEVVLAINVRP 60  
DB 1 MKGLILVGGGTRLRPLTLSPKPLVDFANKPMLHIOEALKEVGVTEVVLAINVRP 60  
QY 61 INFLKDFEDKLGITITCSQETPLCTAGPLALARDKLADSGQDFVFLNSDVISVEYFAE 120  
DB 61 VETLKKEKEYGYNITVSVEYTEPLGTAGPLKLAEDVLKK-DNSPFFVLNSDVICEYFKE 119  
QY 121 LIKPHKHCHGEATIMTKVDPSKYGVVVMEEAT-GRVERVEKPKIFVGNKINAGIYLL 179  
DB 121 LIKPHKHCHGEATIMTKVDPSKYGVVVMEEAT-GRVERVEKPKIFVGNKINAGIYLL 179  
QY 120 LADFHAKHGKGTIVATKVDPSKYGVIVHDIAIPLNIDRVEKPKFVGNRINAGIYLL 179  
DB 120 LADFHAKHGKGTIVATKVDPSKYGVIVHDIAIPLNIDRVEKPKFVGNRINAGIYLL 179  
QY 180 NPSVLDRIELRPYSIEKEVFPQTAADQOLYAMVLPFGVMDVGPDRDITGLRLYLSIRK 239  
DB 180 NPSVLDRIELRPYSIEKEVFPQTAADQOLYAMVLPFGVMDVGPDRDITGLRLYLSIRK 239  
QY 180 NPEVIDLIENKPTSIETETPILVEERQOLYSLFDEGFMMDVGPDKDFLSGTVLYLSLAK 239  
DB 180 NPEVIDLIENKPTSIETETPILVEERQOLYSLFDEGFMMDVGPDKDFLSGTVLYLSLAK 239  
QY 240 KSAAKLATGARHVGNVLVHESAKIGECGLIGPDVAIGPGCVVEDGVRLSRTVMRGVRIK 299  
DB 240 KSAAKLATGARHVGNVLVHESAKIGECGLIGPDVAIGPGCVVEDGVRLSRTVMRGVRIK 299  
QY 240 RQPKKATGANIVGNALIDTAKISSSTAKIGPDVVIGVNTIGDVKILRSVVLNCSSTIK 299  
DB 240 RQPKKATGANIVGNALIDTAKISSSTAKIGPDVVIGVNTIGDVKILRSVVLNCSSTIK 299  
QY 300 KHACISNIIHWSVGVOMARIENMTILGEDVHVHCVDEYVSGVYLPKHKEIKSSILKPEI 359  
DB 300 KHACISNIIHWSVGVOMARIENMTILGEDVHVHCVDEYVSGVYLPKHKEIKSSILKPEI 359  
QY 360 VM 361  
DB 360 IM 361

## RESULT 2

US-09-032-372-3  
Sequence 3, Application US/09032372  
Patent No. 6008337  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Corley, Neil C.  
APPLICANT: Guegler, Karl J.  
APPLICANT: Yue, Henry  
APPLICANT: Lal, Preeti  
TITLE OF INVENTION: CELL CYCLE RELATED PROTEINS  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/032,372  
FILING DATE: Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0478 US  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166

## TELEX:

INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 421 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: COLANOT23  
CLONE: 1693222  
US-09-032-372-3

Query Match 21.5%; Score 401; DB 3; Length 421;

Best Local Similarity 29.1%; Pred. NO. 4.5e-36;  
Matches 123; Conservative 61; Mismatches 159; Indels 80; Gaps 13;

QY 1 MKAILVGG--FGRLRPLTLSPKPLVDFANKPMLHIOEALKEV-GVTEVVLAINVRP 57  
DB 2 LKAVILCGPKQKQKFRPLSPFVPRKPLFPVAGVPMQIHIEACAQVPMQEIILLIGFYQP 61  
QY 58 -EVMINFLKDFEDKLGITITCSQETPLCTAGPLALARDKLADSGQDFVFLNSDVISVEY 116  
DB 62 DEPLTQFLFAAQEENLFVRYLQEFAPLGTGGGLYHFRQILLAGSPEAFVFLNADVCSD 121  
QY 117 PPAELIKPHKHCHGEATIMTKVD--EPSKYGVVVMEEATGRVERVEKPKIFVGNKINA 174  
DB 122 PLSAMLEARRORHPFLLLGTANTQSIINVCIVENPOTHEVLHVVEKPSPTFISDIINC 181  
QY 175 GYLLNPSVL-----DRIELRP-----TSIEKEVFPQTAADQOLYAMV 213  
DB 182 GYLLNPSVL-----DRIELRP-----TSIEKEVFPQTAADQOLYAMV 213  
QY 214 PGFWMVDGPRDYITGLRLYLSIRKSAALATGAHVVGNVL-----VHESAKIGECGL 268  
DB 242 DGIRSQIKSAGSALYASRLYLSRYQDTHPERLA--KHTPGGMDPRECYHPPDRG---- 295  
QY 269 ICPDVAIGP-----GCVVEDGVRLSRL--TVMRGVRIKHKACISNIIHWSV 316  
DB 296 -GPLGCGAPQRLHREGGDRGRCVCGSSGRASSMEPLCRSTRV---FCIASWAGGAPWDAG 351  
QY 317 QWARIENMTI-----LGEDVHVHCVDEYVSGVYLPKHKEI 350  
DB 352 RAWRVPPVTLPTTPEPAWTVRASSRTGSCCLLSPSMCGRCVRIPAEVLILNSIVLPKHKEI 411  
QY 351 KSS 353  
DB 412 SRS 414

## RESULT 3

US-08-196-218-34  
Sequence 34, Application US/08196218  
Patent No. 5614619  
GENERAL INFORMATION:  
APPLICANT: Piepersberg, Wolfgang  
APPLICANT: Stockmann, Michael  
APPLICANT: Taleghani, Kamiz Mansouri  
APPLICANT: Distler, Jurgen  
APPLICANT: Grabley, Susanne  
APPLICANT: Sichel, Petri  
APPLICANT: Brau, Barbara  
TITLE OF INVENTION: Secondary-metabolic Biosynthesis Genes  
FROM ACTINOMYCETES, METHOD OF ISOLATING THEM, AND THEIR  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESS: Dunner  
STREET: 1300 I Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: United States  
ZIP: 20005-3315

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/196,218  
FILING DATE: 25-AUG-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Ogden, Stasia L.  
REGISTRATION NUMBER: 36,228  
REFERENCE/DOCKET NUMBER: 02481.1372-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 355 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-196-218-34

Query Match 16.5%; Score 307; DB 1; Length 355;  
Best Local Similarity 28.4%; Pred. No. 1.1e-25;  
Matches 100; Conservative 59; Mismatches 137; Indels 56; Gaps 12;  
QY 1 MKALILVGGFGTRRLPTLSFPKPLVDFANKPMILHQIEALKEVGVTVE-VLAINYRPEV 59  
Db 1 MKALVLAGGSGTRLRPFSSMPKQLPIANTPVLVHVLNAVRELGVTEGVVGNRGPEI 60  
QY 60 MINLKFDEKLG-----ITTCSETEPLGTAGPLALARDKLADGSGQFFVLNSDVI 113  
Db 61 -----EAVLDGARFDVRIYIPQDAPRGLAHTVSIARGFLGD---DDFVMYLGDNM 109  
QY 114 SEYPPAELIKFKHCHGGEATIMVKVDEPSKYGVVVMEEATGRVERFEKPKIFVGNKIN 173  
Db 110 LPDGVTEIAEEFTQRPAQVVKVPPDRSFGVAEL-GPDGEVLRLVEKWPQPSDMAL 168  
QY 174 AGIYLLNPSVLDR1-ELRPTS-----IEKEVFPQIAADQOLYAMVLPFGWMDVGPDRYI 227  
Db 169 IGYYFTAAIHQAIAAISPSRSGELEITDAVQWLVTSGADVRSALYDGYKDKTGRVEDVL 228  
QY 228 TGLRLYLDLSIRKSAKALATGAHVGNVNLVHESAK-----IGEGCL-----IGP 271  
Db 229 ECNSHLLDGLTPRDVGQDADSVLVGRVYIEAGARIVRSRVEGPAIIGAGTVLQDSQVGP 288  
QY 272 DVAIGPGCVVED-----GVLRSCTV--MRGVRIKKHACISNIIIGHSTVG 316  
Db 289 HTSIGRDTCTVDSRLSGSIALDEASVTGVRGLR-----NSLIGRAASVG 332

RESULT 4  
US-08-681-953-34  
Sequence 34, Application US/08681953  
Patent No. 5710032  
GENERAL INFORMATION:  
APPLICANT: Piepersberg, Wolfgang  
APPLICANT: Stockmann, Michael  
APPLICANT: Taleghani, Kampiz Mansouri  
APPLICANT: Distler, Jurgen  
APPLICANT: Grabley, Susanne  
APPLICANT: Sichel, Petra  
APPLICANT: Brau, Barbara  
TITLE OF INVENTION: Secondary-Metabolite Biosynthesis Genes  
TITLE OF INVENTION: From Actinomycetes, Method of Isolating Them, and Their  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESSEE: Dunner  
STREET: 1300 I Street, N.W.

CITY: Washington  
STATE: D.C.  
COUNTRY: United States  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/681,953  
FILING DATE: 30-JUL-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/196,218  
FILING DATE: 25-AUG-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Ogden, Stasia L.  
REGISTRATION NUMBER: 36,228  
REFERENCE/DOCKET NUMBER: 02481.1372-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 355 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-681-953-34

Query Match 16.5%; Score 307; DB 1; Length 355;  
Best Local Similarity 28.4%; Pred. No. 1.1e-25;  
Matches 100; Conservative 59; Mismatches 137; Indels 56; Gaps 12;  
QY 1 MKALILVGGFGTRRLPTLSFPKPLVDFANKPMILHQIEALKEVGVTVE-VLAINYRPEV 59  
Db 1 MKALVLAGGSGTRLRPFSSMPKQLPIANTPVLVHVLNAVRELGVTEGVVGNRGPEI 60  
QY 60 MINLKFDEKLG-----ITTCSETEPLGTAGPLALARDKLADGSGQFFVLNSDVI 113  
Db 61 -----EAVLDGARFDVRIYIPQDAPRGLAHTVSIARGFLGD---DDFVMYLGDNM 109  
QY 114 SEYPPAELIKFKHCHGGEATIMVKVDEPSKYGVVVMEEATGRVERFEKPKIFVGNKIN 173  
Db 110 LPDGVTEIAEEFTQRPAQVVKVPPDRSFGVAEL-GPDGEVLRLVEKWPQPSDMAL 168  
QY 174 AGIYLLNPSVLDR1-ELRPTS-----IEKEVFPQIAADQOLYAMVLPFGWMDVGPDRYI 227  
Db 169 IGYYFTAAIHQAIAAISPSRSGELEITDAVQWLVTSGADVRSALYDGYKDKTGRVEDVL 228  
QY 228 TGLRLYLDLSIRKSAKALATGAHVGNVNLVHESAK-----IGEGCL-----IGP 271  
Db 229 ECNSHLLDGLTPRDVGQDADSVLVGRVYIEAGARIVRSRVEGPAIIGAGTVLQDSQVGP 288  
QY 272 DVAIGPGCVVED-----GVLRSCTV--MRGVRIKKHACISNIIIGHSTVG 316  
Db 289 HTSIGRDTCTVDSRLSGSIALDEASVTGVRGLR-----NSLIGRAASVG 332

RESULT 5  
US-09-194-905-10  
Sequence 10, Application US/09194905  
Patent No. 6306627  
GENERAL INFORMATION:  
APPLICANT: Decker, Heinrich  
TITLE OF INVENTION: ISOLATION OF THE BIOSYNTHESIS GENES FOR  
TITLE OF INVENTION: PSEUDO-OLIGOSACCHARIDES FROM STREPTOMYCES GLAUCESCENS  
TITLE OF INVENTION: GLA.O AND THEIR USE  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY & LARDNER  
STREET: 3000 K Street, N.W.



CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
FILING DATE: 30-MAY-1997  
APPLICATION NUMBER: US/09/194,905  
FILING DATE: 20-JUL-1998  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/EP97/02826  
FILING DATE: 30-MAY-1997

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 19622783.6  
FILING DATE: 07-JUN-1996

ATTORNEY/AGENT INFORMATION:

NAME: Granados, Patricia D.

REGISTRATION NUMBER: 33,683

REFERENCE/DOCKET NUMBER: 026083/0193

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 672-5300

TELEFAX: (202) 672-5399

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 355 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-194-905-10

Query Match 15.9%; Score 296; DB 4; Length 355;  
Best Local Similarity 27.5%; Pred. No. 1.8e-24;  
Matches 95; Conservative 58; Mismatches 160; Indels 32; Gaps 9;

QY 1 MKALILVGGFGTPLRDLTSLSPKPLVDFANKPMILHQTALKEKVGVTVEVLATN-VRPEV 59  
DB 1 VVALVLGGTGGSKURFTHTAAKOLLPIANKPVLFTALSLAAGVREAGVVGAYGRE- 59  
QY 60 MIFLDFEDKLGITITCSQETPLTAGPLALARDKADGSGQPFVFLNSDVISEYFFA 119  
DB 60 -IRELTGCDTAFGLRTYTHLPRLGLAHAVRIARGPLGD--DDFLYLGDNVLPQGV 115  
QY 120 EIKFKHCHGGEATIMVKVDEPSKYGVVMEERATGRVERFEKPKTFVGNKINAGIYLL 179  
DB 116 DPARQSAADPAARLLLTVPADPSAFGEV-DADGNVLRLEKPDVFRSLALIGVYAF 174  
QY 180 NPSVLDRIE-LRPTS-----IEKEVFPDIAADOLYAMVLPGFMDVQCPDRDVTGLRLY 233  
DB 175 SPAVHEAVRAITPSARGELEITHAVOMMDRGLURVAETTTTRPWRDTGSAEDMLEVNRHV 234  
QY 234 LDISIRKSAK-----LATGAHVGNVNLVHESAKTGCGLIGPDVAIG 276  
DB 235 LDGLEGRIGKVDHSTLVGRVRAECAIVRGSHVGPVIGACAVVSNSS-VGPTYSIG 293  
QY 277 PGCVYEDGVRSLRCTVMRGVRIKKHACISNIIHWHSTYGVQWARI 321  
DB 294 EDCRVDSA-IEYSVLLRGAQVEGASRIEASLIGRGAVVGAPRL 337

## RESULT 6

US-09-134-001C-3316  
Sequence 3316, Application US/09134001C  
Patent No. 6380370

GENERAL INFORMATION:

APPLICANT: LYNN Doucette-Stamm et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: GTC-007

CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-09  
PRIOR APPLICATION NUMBER: US 60/035,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 3316  
LENGTH: 461  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-3316

Query Match 13.0%; Score 241.5; DB 4; Length 461;

Best Local Similarity 25.7%; Pred. No. 3.3e-18;

Matches 98; Conservative 66; Mismatches 159; Indels 59; Gaps 15;

QY 3 ALILVGGFGTPLRDLTSLSPKPLVDFANKPMILHQTALKEKVGVTVEVLAINVRPEVMIN 62  
DB 15 AIIILAAAGGTMMKSKY---KVLHEVAGKPMVHVNKNVKGAGVDQIIVIIHGAE----- 67  
QY 63 FLKDFEDKLGITITCSQETPLTAGPLALARDKADGSGQPFVFLNSDVISEYFFAEL 121  
DB 68 ---SVKDTLGNOSLYSPDKOLCTAHAVKMAHEHLADKECTILVVGCDPLITVOTLOS 124  
QY 122 IKFKHCHGGEATIMVKVDEPSKYGVVMEERATGRVERFEKPKIFVGNK-----INAGIY 177  
DB 125 IEHHESTQSHVTVLSASTINPYGRIIRNH-NGILERIVEEKDANDSEAIKEISSGIF 183  
QY 178 LLNPSVL-DRIELRPTSTIEKEVFPDIAADOLYAMVLPGFMDVQCPDRDVT-----G 229  
DB 184 AFNRVLFKLE-----QVKNDNAGGEYLLDFVLSLILKDGKAKVYCTEDFDEIIG 235  
QY 230 L--RLYLDLSIRKSAK-----LATGAHV--GNVLHESAKTGCGLIGPDVAIGPCV 280  
DB 236 VNDRLMLSEAKALQORINRYHMENGVTIIDPSSTFICTDVKIGIDTIEPCVRIGGHTT 295  
QY 281 YEDGV-----RLSKCTVMRGVRIKKHACISNIIHWHSTYGVQWARIENMTILGEDVHVC 334  
DB 296 IEEDVWTGQVSEINNSTHNSANIKQ-SVTNDSIVGENTVCPAOLRPGSNLASEVKV- 353  
QY 335 DEVYSGGVLPHPKEIKSSILK 356  
DB 354 -----GNFV----EYRKADIK 365

## RESULT 7

US-09-370-838-190  
Sequence 190, Application US/09370838  
Patent No. 6444425  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Lodes, Michael J.  
APPLICANT: Mohamath, Roadon  
APPLICANT: Secrist, Heather  
TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF  
TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE  
FILE REFERENCE: 210121.475C1  
CURRENT APPLICATION NUMBER: US/09/370,838  
CURRENT FILING DATE: 1999-08-09  
EARLIER APPLICATION NUMBER: US 09/285,323  
EARLIER FILING DATE: 1999-04-02  
NUMBER OF SEQ ID NOS: 289  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 190  
LENGTH: 146  
TYPE: PRT  
ORGANISM: Homo sapien  
US-09-370-838-190

Query Match 12.8%; Score 239; DB 4; Length 146;

Best Local Similarity 35.9%; Pred. No. 9.8e-19;

Matches 47; Conservative 23; Mismatches 33; Indels 28; Gaps 1;



4

Db 197 ANPPMPNDPSKSLASMGIVYFDADYLYELLEDDRDNSHDFGKDLIPKITEAGLAYA 256  
QY 211 MVLP-----GFMDVGPDRDITG-----LRLYLDSIRKK 240  
Db 257 HPFLSCVQSDPDAPYWRDVGTLWAYKANLDLASVVPPELMDYRNPRTYNESLPPA 316  
QY 241 SAAKATGAHVGNVNLVHESAKIGEGCLIGPDVAIGPCGVVEDGVRLSCTVMRGVRIKK 300  
Db 317 KFDQDRSGSH--GMTL---NSLVSDGCVIS-----GSVVQSVLFSR-----VRVNS 358  
QY 301 HACISNSIIGHSTVQWARIENMTI-----LGEDVHVCDEVY--SNGGVVLPH 347  
Db 359 FCNIDSALLPEVWVGRSCLRRRCVDRACVPIPEGMVIGENAEEDARRFYRSEEGIVLVT 418  
QY 348 KEI 350  
Db 419 REM 421

RESULT 12  
US-08-398-627-4  
; Sequence 4, Application US/08398627  
; Patent No. 5608149  
; GENERAL INFORMATION:  
; APPLICANT: Barry, Gerard F.  
; APPLICANT: Kishore, Ganesh M.  
; APPLICANT: Stark, David M.  
; TITLE OF INVENTION: Enhanced Starch Biosynthesis  
; NUMBER OF SEQUENCES: 51  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Grace L. Bonner, Monsanto Co. BB4F  
; STREET: 700 Chesterfield Parkway No. 5608149th  
; CITY: St. Louis  
; STATE: Missouri  
; COUNTRY: USA  
; ZIP: 63198  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/398,627  
; FILING DATE: 03-MAR-1995  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/090,523  
; FILING DATE: 12-JUL-1993  
; APPLICATION NUMBER: US 07/709663  
; FILING DATE: 07-JUN-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/53763  
; FILING DATE: 18-JUN-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bonner, Grace L.  
; REGISTRATION NUMBER: 32,963  
; REFERENCE/DOCKET NUMBER: 38-21(10559)A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (314) 537-7286  
; TELEFAX: (314) 537-6047  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 431 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-398-627-4

Query Match 10.2%; Score 191; DB 1; Length 431;  
Best Local Similarity 22.5%; Pred. No. 1.3e-12;  
Matches 95; Conservative 60; Mismatches 170; Indels 98; Gaps 15;

QY 3 ALLIVGGFGTFLRLPLTISFPKPLVDFAKPKMIL-HQIEALKEVGVTEVVLAINRPEVMI 61

Db 22 ALILAGGRTRLKDLTKRAKPAVHFGGKFRIIDFALSNCSINRRMGVITQYQSHTLV 81  
QY 62 NFLK-----DFEDKLGITITCSQETPLGTAGPLALARDKLADSGGQFFVILNS 110  
Db 82 QHIQRGWSFFNEMNEFVLLPAQQRMKGENWYRGTAQVTONLDIIRRYKAEYVILAG 141  
QY 111 DVISEYPPFAELIKFKHCHGGEATI--MVTKVDEPSKYGVVVMEEATGRVERFVEKPKIFV 168  
Db 142 DHIYKODYSRMLIDHVEKGVRCVVCMPVPIEASAFGMVAVDENDKTIIE-FVEKP----- 196  
QY 169 GNK-----INAGIYLLNPSVL-----DRIELRPTSIKEVFPQIAADQOLYA 210  
Db 197 ANPPMPNDPSKSLASMGIVYFDADYLYELLEDDRDNSHDFGKDLIPKITEAGLAYA 256  
QY 211 MVLP-----GFMDVGPDRDITG-----LRLYLDSIRKK 240  
Db 257 HPFLSCVQSDPDAPYWRDVGTLWAYKANLDLASVVPPELMDYRNPRTYNESLPPA 316  
QY 241 SAAKATGAHVGNVNLVHESAKIGEGCLIGPDVAIGPCGVVEDGVRLSCTVMRGVRIKK 300  
Db 317 KFDQDRSGSH--GMTL---NSLVSDGCVIS-----GSVVQSVLFSR-----VRVNS 358  
QY 301 HACISNSIIGHSTVQWARIENMTI-----LGEDVHVCDEVY--SNGGVVLPH 347  
Db 359 FCNIDSALLPEVWVGRSCLRRRCVDRACVPIPEGMVIGENAEEDARRFYRSEEGIVLVT 418  
QY 348 KEI 350  
Db 419 REM 421

RESULT 13  
US-08-406-858-4  
; Sequence 4, Application US/08406858  
; Patent No. 5648249  
; GENERAL INFORMATION:  
; APPLICANT: Barry, Gerard F.  
; APPLICANT: Kishore, Ganesh M.  
; APPLICANT: Stark, David M.  
; APPLICANT: Zalewski, James C.  
; TITLE OF INVENTION: Method of Improving the Quality of  
; TITLE OF INVENTION: Stored Potatoes  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Grace L. Bonner, Monsanto Company, BB4F  
; STREET: 700 Chesterfield Parkway No. 5648249th  
; CITY: St. Louis  
; STATE: Missouri  
; COUNTRY: USA  
; ZIP: 63198  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/406,858  
; FILING DATE:  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/05275  
; FILING DATE: 18-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/070,155  
; FILING DATE: 28-MAY-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bonner, Grace L.  
; REGISTRATION NUMBER: 32,963  
; REFERENCE/DOCKET NUMBER: 38-21(10654)A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (314)537-7286  
; TELEFAX: (314)537-6047

INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 431 AMINO ACIDS  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 DS-08-406-858-4

Query Match 10.2%; Score 191; DB 1; Length 431;  
 Best Local Similarity 22.5%; Pred. No. 1.3e-12;  
 Matches 95; Conservative 60; Mismatches 170; Indels 96; Gaps 15;

QY 3 ALLVCCFCGTRLPDLTSLFPPKDLVDPAKDMIL-HQIEALKEVGVVTVVLAIVRDEVM 61  
 DB 22 AULLAGGRGTRKLDJTNKRKPAVHFGGKFRILDFALSNCINSIGRMGVITQYOSHTLV 81  
 QY 62 NPLK-----DFEDKLGITTCSTQTEPLGTAGPLALARDKLDGSGQPFVVLNS 110  
 DB 92 QHIOGWSFFNEEMNEFVLLPAOORMKGENWYRGTDVATQNLDIRRYKAEYVVLAC 141  
 QY 111 DVISEYFAELIKFKHKGGEATI--MVTKVDEPSKYGVVVMEEATGRVERFVEKPKIFV 168  
 DB 142 DHYKQDYGRMLDHYVEKGVCTVVCMPVIEEASAFGVMAVDENDKTIE-FVEKP----- 196  
 QY 169 GNK-----INAGIYLLNPSVL-----DRIELDPTGIEKEVFPQIAADOOIYA 210  
 DB 197 ANPPSPNDPSKSLASMGIVFDADLYLELLEDDRDENSNDHDFGKDLIPKITEAGLAYA 236  
 QY 211 MVLPL-----GFMDVQGPDDYITG-----LRLYLDLSIRKK 240  
 DB 257 HPFPLSCVQSDDAEPYWRDVTGLAYWKANLDLASVVPDLMDYDRNWPITRYNESLPPA 316  
 QY 241 SAAKLTATGAHVGNVLYHESAKIGEGCLIGPDVAIGPGCVVDEGVRLSCTVMKGVRIKK 300  
 DB 317 KFDVDRSGSH--GMFL--NSLVSQCVLS-----GSVVQSVLFSR-----VRVNS 358  
 QY 301 HACISNIIIGWHSITVCGWARIENMTI-----LGEDVHVCDDEVY--SNGGVVLPH 347  
 DB 359 FCNIDSALLPEVWVGRSCLRCYIDRACVIEGVMVIGENAEEDARRFYRSEEGIVLVT 418  
 QY 348 KEI 350  
 DB 419 REM 421

## RESULT 14

PCT-US91-04036-4

Sequence 4, Application PC/TUS9104036  
 GENERAL INFORMATION:  
 APPLICANT: Kishore, Ganesh M.  
 TITLE OF INVENTION: Increased Starch Content In Plants  
 NUMBER OF SEQUENCES: 23  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Monsanto Co.  
 STREET: 700 Chesterfield Village Parkway  
 CITY: St. Louis  
 STATE: Missouri  
 COUNTRY: USA  
 ZIP: 63198  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US91/04036  
 FILING DATE: 19910607  
 CLASSIFICATION: 800  
 ATTORNEY/AGENT INFORMATION:  
 NAME: McBride, Thomas P.  
 REGISTRATION NUMBER: 32706  
 REFERENCE/POCKET NUMBER: 38-21(10530)A  
 TELECOMMUNICATION INFORMATION:

TELEPHONE: (314) 537-7357  
 TELEFAX: (314) 537-6047  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 431 AMINO ACIDS  
 TYPE: AMINO ACID  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 PCT-US91-04036-4

Query Match 10.2%; Score 191; DB 5; Length 431;  
 Best Local Similarity 22.5%; Pred. No. 1.3e-12;  
 Matches 95; Conservative 60; Mismatches 170; Indels 98; Gaps 15;

QY 3 ALLVCCFCGTRLPDLTSLFPPKDLVDPAKDMIL-HQIEALKEVGVVTVVLAIVRDEVM 61  
 DB 22 AULLAGGRGTRKLDJTNKRKPAVHFGGKFRILDFALSNCINSIGRMGVITQYOSHTLV 81  
 QY 62 NPLK-----DFEDKLGITTCSTQTEPLGTAGPLALARDKLDGSGQPFVVLNS 110  
 DB 82 QHIOGWSFFNEEMNEFVLLPAOORMKGENWYRGTDVATQNLDIRRYKAEYVVLAC 141  
 QY 111 DVISEYFAELIKFKHKGGEATI--MVTKVDEPSKYGVVVMEEATGRVERFVEKPKIFV 168  
 DB 142 DHYKQDYGRMLDHYVEKGVCTVVCMPVIEEASAFGVMAVDENDKTIE-FVEKP----- 196  
 QY 169 GNK-----INAGIYLLNPSVL-----DRIELRPTSIEKEVFPQIAADOOIYA 210  
 DB 197 ANPPSPNDPSKSLASMGIVFDADLYLELLEDDRDENSNDHDFGKDLIPKITEAGLAYA 256  
 QY 211 MVLPL-----GFMDVQGPDDYITG-----LRLYLDLSIRKK 240  
 DB 257 HPFPLSCVQSDDAEPYWRDVTGLAYWKANLDLASVVPDLMDYDRNWPITRYNESLPPA 316  
 QY 241 SAAKLTATGAHVGNVLYHESAKIGEGCLIGPDVAIGPGCVVDEGVRLSCTVMKGVRIKK 300  
 DB 317 KFDVDRSGSH--GMFL--NSLVSQCVLS-----GSVVQSVLFSR-----VRVNS 358  
 QY 301 HACISNIIIGWHSITVCGWARIENMTI-----LGEDVHVCDDEVY--SNGGVVLPH 347  
 DB 359 FCNIDSALLPEVWVGRSCLRCYIDRACVIEGVMVIGENAEEDARRFYRSEEGIVLVT 418  
 QY 348 KEI 350  
 DB 419 REM 421

## RESULT 15

PCT-US94-05275-4

Sequence 4, Application PC/TUS9405275  
 GENERAL INFORMATION:  
 APPLICANT:  
 TITLE OF INVENTION: Method of Improving the Quality of Stored  
 NUMBER OF SEQUENCES: 26  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US94/05275  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/070155  
 FILING DATE: 28-MAY-1993  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 431 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein

PCT-US94-05275-4

Query Match		10.2%;	Score 191;	DB 5;	Length 431;
Best Local Similarity		22.5%;	Pred. No. 1.3e-12;		
Matches	95;	Conservative	60;	Mismatches	170;
				Indels	98;
				Gaps	15;
QY	3	ALILVGGFRLRLPLTSLSPKPLVDFANKPMIL-HQTEALKEVGVTEVVLAINYRPEVMI	61		
Db	22	ALILAGRGTRKDLTNKRAKPAVHFGGKFRIIDFALSNCINSIGRRMGVITQYOSHTLV	81		
QY	62	NFLK-----DFEDKLGITITCSQETEPGLTAGPLALARDKLADSGGQFFVLNS	110		
Db	82	QHIQRGWSFENEMNEFVDLLPAQRKMGENWYRGTAQVNTQNLDIIRRYKAQYVVILAG	141		
QY	111	DVISEYPPFABLKFHKHCGEATI--MVTKVDEPSKYGVVMEATGRVERFEVEKPIFV	168		
Db	142	DHIYKQDYSRMLDHDHVEKGYRCHVVCMPVPIEASAFGVMAVDENDKTIE-FVEKP----	196		
QY	169	GNK-----INAGIYLLNPVL-----DRIELRPTSIEKEVFPQIAADQOLYA	210		
Db	197	ANPPSPNDPSKSLASNGIYVFDADYLYELLEEDDRDENSDDFGKDLIPKITEAGLAYA	256		
QY	211	MVLP-----GFWMVVGQPRDYITG-----LRLYLDSIRKK	240		
Db	257	HPFPLSCVQSDPDPAEPYWRDVGTLFAYWKANLDIASVVPFLDMYDRNWPRTYNESLPPA	316		
QY	241	SAAKLATGAHVGNVLVHESAKIGEGCLIGPDVAIGPCVVEDGCVRLSRCVTVMRGVRIKK	300		
Db	317	KFQDRSGSH--GWTL---NSLVSDGCVIS-----GSVVQSVLFSR-----VRVNS	358		
QY	301	HACISNSIIGWHSVTVGQWARIENNTI-----LGEDVHVCDDEVY--SNGGVVLPH	347		
Db	359	FCNIDSAVLLPEYVWVGESCLRLRCVIDRACVIEGHWIGENABEDARRFYRSEEGIVLVT	418		
QY	348	KEI	350		
Db	419	REM	421		

Search completed: November 27, 2002, 05:11:25  
Job time : 28 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 05:09:42 ; Search time 38 seconds  
(without alignments)  
151.280 Million cell updates/sec

Title: US-09-374-967-2  
Perfect score: 1864  
Sequence: 1 MKALLVGGFGTRLRLTLTSLT.....GVVLPKHKESILKPEIVM 361

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 102317 seqs, 15924203 residues

Total number of hits satisfying chosen parameters: 102317

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published\_Applications\_AA:\*
- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pap:\*
  - 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pap:\*
  - 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pap:\*
  - 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pap:\*
  - 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pap:\*
  - 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pap:\*
  - 7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pap:\*
  - 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pap:\*
  - 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pap:\*
  - 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pap:\*
  - 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pap:\*
  - 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pap:\*
  - 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pap:\*
  - 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	488	26.2	359	9	US-09-712-363-269
2	277.5	14.9	433	10	US-09-925-637-18
3	275.5	14.8	449	10	US-09-815-242-5479
4	275.5	14.8	452	10	US-09-815-242-12149
5	239	12.8	146	10	US-09-738-973-190
6	227	12.2	456	10	US-09-815-242-10400
7	216.5	11.6	458	10	US-09-815-242-11759
8	216	11.6	456	10	US-09-815-242-14064
9	209	11.2	456	10	US-09-815-242-11073
10	190	10.2	292	9	US-09-860-846-12
11	190	10.2	292	10	US-09-861-289-12
12	190	10.2	3782	9	US-09-860-846-4
13	190	10.2	3782	10	US-09-861-289-4
14	189.5	10.2	471	10	US-09-815-242-10937
15	185	9.9	462	10	US-09-815-242-13542
16	185	9.9	479	10	US-09-815-242-13307
17	179	9.6	454	10	US-09-815-242-12112
18	174	9.3	440	10	US-09-734-569-178
19	162.5	8.7	293	10	US-09-934-868-22

20	162.5	8.7	293	10	US-09-934-868-22	Sequence 22, Appl
21	142	7.6	433	10	US-09-815-242-11355	Sequence 11355, A
22	142	7.6	445	10	US-09-881-752A-274	Sequence 274, App
23	135.5	7.3	160	10	US-09-738-973-189	Sequence 189, App
24	120	6.4	237	10	US-09-765-272-136	Sequence 136, App
25	94	5.0	398	10	US-09-896-852-54	Sequence 54, Appl
26	94	5.0	506	10	US-09-896-852-52	Sequence 52, Appl
27	94	5.0	551	10	US-09-896-852-53	Sequence 53, Appl
28	94	5.0	667	10	US-09-896-852-55	Sequence 55, Appl
29	93.5	5.0	236	10	US-09-815-242-5707	Sequence 5707, Ap
30	93.5	5.0	242	10	US-09-815-242-12467	Sequence 12467, A
31	92	4.9	704	10	US-09-912-020-248	Sequence 248, App
32	92	4.9	704	10	US-09-815-242-10363	Sequence 10363, A
33	91	4.9	171	10	US-09-815-242-11575	Sequence 11575, A
34	90	4.8	238	10	US-09-876-182-2	Sequence 2, Appli
35	89.5	4.8	360	10	US-09-841-132-506	Sequence 506, App
36	88.5	4.7	388	10	US-09-938-540-2	Sequence 2, Appli
37	88	4.7	449	10	US-09-925-301-1299	Sequence 1299, Ap
38	87	4.7	704	10	US-09-815-242-13965	Sequence 13965, A
39	86.5	4.6	1981	9	US-09-928-457-38	Sequence 38, Appl
40	85.5	4.6	354	10	US-09-841-132-576	Sequence 576, App
41	84.5	4.5	608	10	US-09-767-041-39	Sequence 39, Appl
42	84	4.5	57	10	US-09-734-017A-52	Sequence 52, Appl
43	84	4.5	182	10	US-09-734-569-4	Sequence 4, Appli
44	84	4.5	184	10	US-09-734-569-144	Sequence 144, App
45	84	4.5	1252	10	US-09-841-132-449	Sequence 449, App

ALIGNMENTS

RESULT 1

- US-09-712-363-269
- ; Sequence 269, Application US/09712363
  - ; Patent No. US20020164588A1
  - ; GENERAL INFORMATION:
  - ; APPLICANT: Eisenberg, David
  - ; APPLICANT: Rotstein, Sergio H.
  - ; APPLICANT: Marcotte, Edward M.
  - ; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS
  - ; FILE REFERENCE: 07419-032001
  - ; CURRENT APPLICATION NUMBER: US/09/712,363
  - ; CURRENT FILING DATE: 2000-11-13
  - ; PRIOR APPLICATION NUMBER: PCT/US00/02246
  - ; PRIOR FILING DATE: 2000-01-28
  - ; PRIOR APPLICATION NUMBER: 60/179,531
  - ; PRIOR FILING DATE: 2000-02-01
  - ; PRIOR APPLICATION NUMBER: 60/117,844
  - ; PRIOR FILING DATE: 1999-01-29
  - ; PRIOR APPLICATION NUMBER: 60/118,206,
  - ; PRIOR FILING DATE: 1999-02-01
  - ; PRIOR APPLICATION NUMBER: 60/126,593
  - ; PRIOR FILING DATE: 1999-03-26
  - ; PRIOR APPLICATION NUMBER: 60/134,093
  - ; PRIOR FILING DATE: 1999-05-14
  - ; PRIOR APPLICATION NUMBER: 60/134,092
  - ; PRIOR FILING DATE: 1999-05-14
  - ; PRIOR APPLICATION NUMBER: 60/165,124
  - ; PRIOR FILING DATE: 1999-11-12
  - ; PRIOR APPLICATION NUMBER: 60/165,086
  - ; PRIOR FILING DATE: 1999-11-12
  - ; NUMBER OF SEQ ID NOS: 292
  - ; SOFTWARE: FastSeq for Windows Version 4.0
  - ; SEQ ID NO 269
  - ; LENGTH: 359
  - ; TYPE: PRT
  - ; ORGANISM: Mycobacterium tuberculosis
  - ; US-09-712-363-269

Query Match 26.2%; Score 488; DB 9; Length 359;  
Best Local Similarity 34.6%; Pred. No. 1.1e+40;  
Matches 125; Conservative 63; Mismatches 151; Indels 22; Gaps 10;

```
QY 3 ALLIVGGFCTRLRPLTLSPKPLVDFAKPMILHQLKALKEVCVTEVWLAINRPEVMIN 62
DB 8 AVVLYGGKGTRLRPLTLSPKPLVDFAKPMILHQLKALKEVCVTEVWLAINRPEVMIN 64
QY 63 FLKDFED--KLGIITTCSETEPLCTAGPLALARDKLDGSCGPPFFVLSNDVISEVPAEL 120
DB 65 FRAEFGDSALGLQIEYVTEHPGLTGGGIANVAGLRNDTA---WVFGDVLUGADLAQ 121
QY 121 LKFKHCHGGEATIMVTKVDEPSKYGVVYVMEATGRVERFEVKFIYFGKKNINAGIYLLN 180
DB 122 LLDHRSNRADVTQLVRVGDPRAFGCVPTDE-EDRVVAFLEKTEDDPDPTQINAGCVVPE 180
QY 161 PVLDRIT-ELRSTSEKVEFPOIAAQ--OLYAMVLPFGFWMQVGPQRYDITGLRLYLDISI 237
DB 181 RNVIDRIPQGREVSVERVEFPALLADGCKIKGYVDASVWRDMGTDEDFVRG---SADLV 237
QY 238 RKKSAKLATGAHVGNVLAVHESAKIGEGCLICPDVAIGPCCVVEDGVRLSRCTVMRGVR 297
DB 238 RGIAFSPALRGHR--GEQJYHDAAYSPGALLIGGTVVGGAIEGPGTRLDGGAVIDGVR 295
QY 298 IKKHACISNIIHSTVGVQWARIENNTI-LGEDVHVCDVYSNG---GVVLPKHEIKS 352
DB 296 VEAGCVIERSIIGFGARIGPRALIRGVICGDGADICARCELLSGARVWPGVFLPDGGIRY 355
QY 353 S 353
DB 356 S 356

RESULT 2
us-09-925-637-18
; Sequence 18, Application us/09925637
; Patent No. US2002010338A1
; GENERAL INFORMATION:
; APPLICANT: CHOI
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Polypeptides
; FILE REFERENCE: PB560
; CURRENT APPLICATION NUMBER: US/09/925,637
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/23773
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 60/151,933
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: US 08/781,986
; PRIOR FILING DATE: 1997-01-03
; PRIOR APPLICATION NUMBER: US 08/956,171
; PRIOR FILING DATE: 1997-10-20
; PRIOR APPLICATION NUMBER: US 60/009,861
; PRIOR FILING DATE: 1996-01-06
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Homo sapiens
us-09-925-637-18
```

Query Match 14.9%; Score 277.5; DB 10; Length 453;  
Best Local Similarity 27.2%; Pred. No. 1.2e-19;  
Matches 106; Conservative 64; Mismatches 146; Indels 73; Caps 15;

```
QY 3 ALLIVGGFCTRLRPLTLSPKPLVDFAKPMILHQLKALKEVCVTEVWLAINRPEVMIN 62
DB 8 AVVLYGGKGTRLRPLTLSPKPLVDFAKPMILHQLKALKEVCVTEVWLAINRPEVMIN 64
QY 63 FLKDFEDKLGIITTCSETEPLCTAGPLALARDKLDGSCGPPFFVLSNDVISEVPAEL 121
DB 61 ---SVKHLGERSLSYFQDEQLGTAHAYOMAKSHLEDKEGTTIVVCGDPTLIKTETVL 117
QY 122 IKFKHCHGGEATIMVTKVDEPSKYGVVYVMEATGRVERFEVKFIYFGKKNINAGIY 177
DB 118 IAHEDANAQATVLSASIQOPYGGRIV-RNASGRLERIVEEKDAQAEKDINEISSGIF 176
```

```
QY 178 LLNPSVLDRIELRPTSTSEKVEFPOIAAQQOQLYAMVLP-----GFWMVDVGQPRDY-- 226
DB 177 AFNNKTL-----FEKLTQVKNDNAOGEYYLDPVLSLILNDGGIVEYRTNDVVEE 225
QY 227 ITGL--RLYLDSTIRKSAKALATGAHVYGNLYVHESAKIGEGCLIGPQYALGPGCVYEDG 284
DB 226 IMGVNDRVML-SQAEKAMORRTNHYHMLNGVTIIDP-----DSTVIGDDVTTCSDVTIEPG 280
QY 285 VRLS-----RCTVMRGVYRIKKHACISNSI-----IGHSTVSGOMARIENMTIL 327
DB 281 VRINGRTEIGEDVVYIGQVSEINNSTIENGACIQGVVNDASVGANTKVGPFPQALRPGAQL 340
QY 328 GEDVHVCDVYSNCGVVLPHKEIKSSILK 356
DB 341 GADYKV-----GNFY-----EIKKADLK 358

RESULT 3
US-09-815-242-5479
; Sequence 5479, Application us/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 5479
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5479
```

Query Match 14.8%; Score 275.5; DB 10; Length 449;  
Best Local Similarity 27.2%; Pred. No. 1.9e-19;  
Matches 106; Conservative 64; Mismatches 146; Indels 73; Caps 15;

```
QY 3 ALLIVGGFCTRLRPLTLSPKPLVDFAKPMILHQLKALKEVCVTEVWLAINRPEVMIN 62
DB 4 AIIAAGKGTMRKSKY---KYLHEVAGKPMVEHLESVKSGVQDVVTIVGHGAE---- 56
QY 63 FLKDFEDKLGIITTCSETEPLCTAGPLALARDKLDGSCGPPFFVLSNDVISEVPAEL 121
DB 57 ---SVKHLGERSLYSPDEQLGTAHAYOMAKSHLEDKEGTTIVVCGDPTLIKTETVL 113
QY 122 IKFKHCHGGEATIMVTKVDEPSKYGVVYVMEATGRVERFEVKFIYFGKKNINAGIY 177
DB 114 IAHEDANAQATVLSASIQOPYGGRIV-RNASGRLERIVEEKDATAEKDINEISSGIF 172
```



Qy	178	LLNPSVLDRIELRPT	SIEKEVFEPQ	IAADQOLYAMVLP	-----GFWMYDQPRDY--	226	
Db	173	AFNNKTL	-----FEKLTQVKNDAQGEY	LPDVL	SLILNDGGIVEYVRNDVBE	221	
Qy	227	ITGL--RLYLD	SRKSAKLATGAHV	GVNVLVHESAK	IGECCLIGPDVAIGPCWVEDG	284	
Db	222	IMGVNDKVM	--SQAEKAMQRRTN	HHMLNGVTIIDP	-----DSTF	IGPDVILGSDVTIEPG	276
Qy	285	VRLS	-----RCVTMRGVR	IKKNACTNSI	-----IGWISTVGOWARIENMTL	327	
Db	277	VRINGRTEIGEDV	VIQYSEINN	STIENGAC	LQQSVVNDASVGANTKVGPF	QAQLRPAQL	336
Qy	328	GEDVHVCDVYSNG	VVLPHKEIKKSILK			356	
Db	337	GADVKV	-----GNEV	-----ETKKADLK		354	

## RESULT 4

```

US-09-815-242-12149
; Sequence 12149, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELTRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12149
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12149

```

```

QY 178 LLNPSVLDRIELRPTSIKEVFPQIAADQQLYAMVLP-----GFMMDVGQPRDY-- 226
      |         |         |         |         |         |         |         |
Db 176 AFNKKTL-----FEKLTQVKNDNAQGEYYLPDVLSLILNDGGIVEYVRTNDVEE 224
      |         |         |         |         |         |         |         |
QY 227 ITGL--RLYLDISIRKSKAKLATGAHVGVNVLVHESAKIGEGCLIGPDVATGPGCVVEDG 284
      | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 225 IMGVNDRVML--SQAEKAMQRRTNHYHMLNGVTIIDP-----DSTFTGPDVITIGSDTVIEPG 279
      | | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 285 VRLS-----RCTVMRGVIRKKHKAICISNSI-----IGHWSTVGGOWARIENMTIL 327
      | | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 280 VRINGRTEIGEDVVGIVQYSEINNSTIENGACIQQSVNDASVGANKTKVGPFAQLRPGAQL 339
      | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 328 GEDVHVCDEVYNSNGGVLPKHEIKSILK 356
      | | | | | | | | | | | | | | | | | | | | | | | | |
Db 340 GADVKV-----GNFV----EIRKADLK 357
      | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 5
US-09-738-973-190
; Sequence 190, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Fling, Steven P.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Algate, Paul A.
; APPLICANT: Secrist, Heather
; APPLICANT: Indirias, Carol Yoseph
; APPLICANT: Benson, Darin R.
; APPLICANT: Elliot, Mark
; APPLICANT: Mannion, Jane
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C9
; CURRENT APPLICATION NUMBER: US/09/738,973
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 190
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-738-973-190

```

## RESULT 6

RES-09-815-242-10400  
 ; Sequence 10400, Application US/09815242  
 ; Patent No. US20020061569A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Haselbeck, Robert  
 ; APPLICANT: Ohlsen, Kari L.  
 ; APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
TITLE OF INVENTION: Prokaryotes  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 10400  
LENGTH: 456  
TYPE: PRT  
ORGANISM: Escherichia coli  
US-09-815-242-10400

Query Match 12.2%; Score 227; DB 10; Length 456;  
Best Local Similarity 24.4%; Pred. No. 13e-14;  
Matches 100; Conservative 55; Mismatches 158; Indels 96; Gaps 16;  
QY 1 MKALILGVGGTGLRPLTLSPFKPLVDFAKPKMLHQIEALKVEGVTEVYVLAIRPEVM 60  
Db 6 MSVVILAAGKGTMYSDLPKVLHTLAGKAMVQHVDAANELAAHVLVYGHGGDLL 62  
QY 61 INFLKDFEDKLGITITCSQETEPGLTAGPLALARDKLDGSGOPFFVLNSDVISEYPPAE 120  
Db 63 KOALKD-DNLNWL---QAEQLGTGHAMQQA-----APFFADDEDILMLYGDVP 107  
QY 121 LKFKH-----KCHGEATIMTKVDPSKYGVVME--EATGRVER--FVEPKIF 167  
Db 108 LLSVETLORLRAKPOGG-IGLLTVKLDPTGYGRITRENGKVTGIVEHKKDAIDEORQI- 165  
QY 168 YGNKINAGIYLLN-----PSVLDRI 187  
Db 166 --QEINTGILIANGADMKRWLAKLTNNNAQGEVYITDIALAVQEGREIVAVHPQLSEV 223  
QY 188 ELRPTSIEKEVPQTAADQOLYAMVLPGFWMVGQPRDYIT-GLRVLDSIRKKSAAKLA 246  
Db 224 EGVNRLQLSLRERYVQSEAEKLLLAGVNL-----RDPAREDLRGTLTHGRD---VEID 275  
QY 247 TCAHVGVNVLVHESAKIGEC-----LIGPDVAIGCCVVEDGVRLSRCTV-----MRG 295  
Db 276 TNNVLEGNVTGHRVKIGTCVKNSVIGDCEISPYTVVEDANLAAACTIGPFARLRP 335  
QY 296 VRIKKHACISNLSIGHVSTVGOWARNTILGEDVHVCDEVYSGNGVV 344  
Db 336 ALLEGAHVGNFVEMKKARLGKSKAGHLYLG-DAEIGNVNIAGTI 383

RESULT 7  
US-09-815-242-11759  
; Sequence 11759, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
TITLE OF INVENTION: Prokaryotes  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 11759  
LENGTH: 458  
TYPE: PRT  
ORGANISM: Klebsiella pneumoniae  
US-09-815-242-11759

Query Match 11.6%; Score 216.5; DB 10; Length 458;  
Best Local Similarity 23.8%; Pred. No. 1.5e-13;  
Matches 100; Conservative 62; Mismatches 156; Indels 105; Gaps 17;  
QY 1 MKALILGVGGTGLRPLTLSPFKPLVDFAKPKMLHQIEALKVEGVTEVYVLAIRPEVM 60  
Db 8 MSVVILAAGKGTMYSDLPKVLHTLAGKPMYQHVDAANDLGAACAVHLVYGHGGDLL 64  
QY 61 INFLKDFEDKLGITITCSQETEPGLTAGPLALARDKLDGSGOPFFVLNSDVISEYPPAE 120  
Db 65 RQTL--HEDNLNWL---QAEQLGTGHAMQQA-----APFFNDEOILMLYGDVP 109  
QY 121 LKFKH-----KCHGEATIMTKVDPSKYGVVME--EATGRVEREVEKPKIFGN 170  
Db 110 LISVETLORLRAKPOGG-IGLLTVKLDPTGYGRITRENGKVTGIVEHKKDAISEAORIO 168  
QY 171 KINAGIYLLN-----LNFSLVDRIE-- 188  
Db 169 EINTGILIANGADLKRWLAKLTNNNAQGEYITDIAAHQEGHQIVAVHPQLSEVGV 228  
QY 189 ---LRPTSIEKEVPQTAADQOLYAMVLPGFWMVGQPRDYITGLRLVLDIRKKSAAKL 245  
Db 229 NNRLQALRLR-VYQAEQAEKLLLAGVNL-----LRDPARF--DLR---GTLQHGROVEI 276  
QY 246 ATGAHVGVNVLVHESAKIGECGL-----IGPDVAIGCCVVEDGVRLSRCTV-----MR 294  
Db 277 DTNVILEGNVTGHRVKIGACGVTKNSTIGDCEISPYSVVEDAQLHAACCTIGPFARLRP 336  
QY 295 GVRIKKHACISNLSIGHVSTVGOWARNTILGEDVHVCDEVYSGNGVV-----PH 347  
Db 337 GAELLEGAHVGNFVEMKKARLGKSKAGHLYLG-DAEIGNVNIAGTTCNVDCAKWH 395  
QY 348 KEI 350  
Db 396 KTI 398  
RESULT 8  
US-09-815-242-14064  
; Sequence 14064, Application US/09815242  
; Patent No. US20020061569A1



[illegible]

```

; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10937
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10937

Query Match          10.2%; Score 189.5; DB 10; Length 461;
Best Local Similarity 21.5%; Pred. No. 7.1e-11;
Matches 87; Conservative 70; Mismatches 158; Indels 89; Gaps

Qy 3 ALILVGGFGRRLRPLTISFPKPLVDFAKPMILHQIBALKEVGVEVVLAINRYRPEVMIN 62
    I:|I| I|I|I| I|I|I| I|I|I| I|I|I| I|I|I| I|I|I| I|I|I| I|I|I|
Db 9 AIIILAAGKGRMKS--KLYKVLHPVAGKPMVEHILDQVEQTEPTIVTIVGHGAEMI-- 63

Qy 63 FLKDFEDKLGITITCSQETPEPLGTAGPALARDKLDGSGQPPFVL-NSDVISPEYPAEL 121
    :|I| :|I| :|I|I| :|I|I| :|I|I| :|I|I| :|I|I| :|I|I| :|I|I|
Db 64 -----KSHLGRS0YALQAEOLGTGHAVMQAQLLGLGKQGTLLVITGDTPLLTAAETLKNL 118

Qy 122 IKPHKCHGGEATIMVTKVDEPSKGVVVVMEATGRVERFEVEKPKIFVG----NKINAGIY 177
    :|I| :|I|I| :|I|I| :|I|I| :|I|I|I| :|I|I|I| :|I|I|I| :|I|I|I|
Db 119 FDYHQGNKASATILTAHAEDPTGCGRIIRDH-VGIVIRIVEQKDASEEARVQEIQTGF 177
    I|I|I|I| :|I| :|I| :|I|I|I|DIIEILKEGKAVAAQY-----MADFDAMGVN 231

Qy 178 LL-NPSVLDRIELRPTS-----IKEVFVQIADQOQLYAMVLPGFWMMDVG-- 221
    I|I|I|I| :|I| :|I| :|I|I|I|DIIEILKEGKAVAAQY-----MADFDAMGVN 231
Db 178 CFDNESLFEALAKTDTNNTQGEYVLTDIIEILKEGKAVAAQY-----MADFDAMGVN 231

Qy 222 -----QPRDYITGL---RLYDSIRKKSAAKATGAHVGVNVLVHE 259
    :|I| :|I| :|I| :|I| :|I| :|I| :|I| :|I| :|I| :|I| :|I| :|I|
Db 232 DRVALSTANKIMHRRLENMHRNGVTFIDPDTYID-----EGVIVGSDTVIEAGVTIKG 286

Qy 260 SAKIGECCLIGP-----DVAIGPCVVVEDGVRLSRCTVMRGVRIKKHACI-SNSIIGWHS 313
    I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|
Db 287 KTVIGEDCLGAHSEIVDSHIGNGVVVKQSV-IEESVVRGADVGPYAHLRPRADVGANV 345

Qy 314 TVGWQARIENNTI-----LGEDVHV-CDEYISN 340
    :|I| :|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|
Db 346 HIGNFVEVKNATIDEGTKVGHLYTVGDATLGKDINVCGVVFN 389

RESULT 15
US-09-815-242-13542
; Sequence 13542, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haseibeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITFA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625

```

```
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/369,308
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 14110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 13542
/ LENGTH: 472
/ TYPE: PRT
/ ORGANISM: Streptococcus pneumoniae
US-09-815-242-13542

Query Match      9.9%; Score 185; DB 10; Length 472;
Best Local Similarity 24.3%; Pred. No. 2,1e-10;
Matches 90; Conservative 56; Mismatches 175; Indels 50; Gaps 11;

QY 3 ALLILGCGFTRLRPLILSPKPLVDFAKPMILHQIEALKEVGYTEVYVLAINRPEVMIN 62
DB 18 AIIAAGKGRKMS---DLFKVLHKVAGISMLEHVFPRSGAIOPEKTVTVVGHKAELV-- 72
QY 63 FLKDPEDKLGITITCSQETPEPLCTAGPLALARDKLDGSGOPFFVL-NSDVISEYPFAEL 121
DB 73 -----EEVLAEQTEFTVTSQGLGIGHAVMTEFIEGLSGHTLVIAAGDTPLITGESLKNL 127
QY 122 IRFKHCHGGEATIMTKVDPSKYGVVVMEEATGRVERFVEKPKIFVGNK-----INAGIY 177
DB 128 IDFHINHKNVATILTAETDNPFYGRIVRND-NAEVLRIVEQKDATDFEKQKEINTGY 186
QY 178 LL-NPSYLDRIELRPTSIEKEVFQIAADQQLYAMVLPGFMDVGCQPRDYITGLRLYLD 236
DB 187 VFDNERLFEALKNINFN-----NAOGEYVITDVIGIFRETGEKVGAYT-LKDFDES 236
QY 237 IRKKSAAKLATGAHVGNVLVHESAKIG-----EGCLIGPDVAIGPGCVVEDGV----- 285
DB 237 LGVNDRYALATAESYRRRINHRHMVNGSVFVPEATYIDIDVEIADEVOIEANVILKQ 296
QY 286 -----RLSRCTVMRGVRIKKHACISNSIICWHSS-----TVGOWARIENMTILGEDVHV 333
DB 297 TKIGAETVLTNGTYVVDSTIGAGAVITNSMIEESSVADGVTVGPIAHIRENSSLGQVHI 356
QY 334 CDEVSYNGGVV 344
DB 357 GNFVEVKGSSI 367
```

Search completed: November 27, 2002, 05:19:00  
JOB time : 42 secs

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	1681	90.2		361	2	T01007	mannose-1-phosphat
2	1587.5	85.2		364	2	T47698	mannose-1-phosphat
3	1376	73.8		351	2	F85357	GDP-mannose pyroph
4	1178	63.2		361	2	S67590	mannose-1-phosphat
5	1175	63.0		363	2	T41209	mannose-1-phosphat
6	1164	62.4		363	2	T42371	probable mannose-1
7	1045	56.1		373	2	T32900	hypothetical prote
8	550	29.5		415	2	G96778	hypothetical prote
9	504	27.0		842	2	AB3101	mannose-1-phosphat
10	493.5	26.5		385	2	G84459	probable GDP-manno
11	492	26.4		831	2	A70363	mannose-1-phosphat
12	490.5	26.3		414	2	T39403	probable mannose-1
13	488	26.2		359	2	A70978	probable rmlA2 pro
14	478	25.6		413	2	H75070	sugar-phosphate nu
15	475	25.5		416	2	A71095	probable sugar-pho
16	473.5	25.4		361	2	F90174	sugar phosphate nu
17	457.5	24.5		843	2	S76785	hypothetical prote
18	451.5	24.2		358	2	B87003	hypothetical prote
19	444	23.8		361	2	G75164	mannose-1-phosphat
20	421.5	22.6		361	2	C71177	probable mannose-1
21	420.5	22.6		385	2	G63101	mannose-1-phosphat
22	398	21.4		815	2	G97266	mannose-1-phosphat
23	392	21.0		423	2	A69079	glucose-1-phosphat
24	377.5	20.3		234	2	A97276	nucleoside-diphosp
25	377.5	20.3		352	2	A75098	glucose-1-phosphat
26	375.5	20.1		395	2	F84160	glucose-1-phosphat
27	370	19.8		353	2	T48521	probable dTDP-1-gl
28	365	19.6		361	2	A75596	probable glucose-1
29	359.5	19.3		356	2	F71151	probable glucose-1

Db 181 PSVLKDIELPRPSIEKETPPKIAAAGCLYAMVLPGFWMIDGQPRDITGLRLYDLSURKK 240  
Qy 241 SAAKLATGAHVGNVNLVHESAKIGECGLICPDVAIGPGCVVEDGVRLSRTVMRGVRIKK 300  
Db 241 SPAKLTSGPHIYGNVLYDETATIGESGLIGPDVAIGPGCVIESGVRLSRTVMRGVRIKK 300  
Qy 301 HACISNSIIGWSTVQGWARIENMTILGEDVHVCDDEVYNGGVLPKHKEKSKILKPEIV 360  
Db 301 HACISSIIIGWSTVQGWARIENMTILGEDVHVSDRIYSNGGVLPKHKEKSKILKPEIV 360  
Qy 361 M 361  
Db 361 M 361

RESULT 2  
T47698  
mannose-1-phosphate guanylyltransferase-like protein - Arabidopsis thaliana  
N:Alternate names: protein T22E16.250  
C:Species: Arabidopsis thaliana (mouse-ear cross)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 19-May-2000  
C:Accession: T47698  
R:Benes, V.; Wurmbach, B.; Drzonek, H.; Ausorge, W.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.  
submitted to the Protein Sequence Database, February 2000  
A:Reference number: Z24472  
A:Accession: T47698  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-364 <BEN>  
A:Cross-references: EMBL:AL132975  
A:Experimental source: cultiwar Columbia; BAC clone T22E16  
C:Genetics:  
A:Map position: 3  
A:Introns: 37/3; 58/3; 141/3  
A:Note: T22E16.250  
C:Superfamily: mannose-1-phosphate guanylyltransferase

Query Match 85.2%; Score 1587.5; DB 2; Length 364;  
Best Local Similarity 83.2%; Pred. No. 2.4e-118;  
Matches 303; Conservative 29; Mismatches 29; Indels 3; Gaps 1;

Qy 1 MKALILVGGCGTRLRPLTSLSPFKPLVDFAFKPMILHQIEALKEVGYTEVVLAINYPE-- 58  
Db 1 MKALILVGGCGTRLRPLTSLPKPLVDFAFKPMILHQIEALKAIGDEVVLAINYPEQL 60  
Qy 59 -VMNFKDFEDKLGITITCSQETPLGTAGPLALARDKLADGSGQFFVNLSDVISEYP 117  
Db 61 LVMKFSNDVEATLGKIKITCSQETPLGTAGPLALARDKLVDGSGQFFVNLSDVISDYP 120  
Qy 118 PAELIKPHKCHGGEATIMVTKVDEPSKGVVWMEETGVRERVEKPIFVGKNINAGIV 177  
Db 121 LEEMTAFNHAHGEASIMVTKVDEPSKGVVWMEETGVRERVEKPKIFVGKNINAGIY 180  
Qy 178 LNPVSLDRIELRPTSIKEIEFPQIAADQOLYAMVLPGFWMIDGQPRDYITGLRLYDLSI 237  
Db 181 LNPVSLDRIELRPTSIKEIEFPQIAAEKLYAMVLPGFWMIDGQPRDYITGLRLYDLS 240  
Qy 238 RKSAKLATGAHVGNVNLVHESAKIGECGLIGPDVAIGPGCVVEDGVRLSRTVMRGVR 297  
Db 241 RKKSPKLTATGPHILGNVLYDETAIEGECGLIGPVAIGPGCVYEGYRLSHCTYMRGVH 300  
Qy 298 IKKHACISNSIIIGWSTVQGWARIENMTILGEDVHVCDDEVYNGGVLPKHKEKSKILKP 357  
Db 301 VARYACISSIIIGWSTVQGWARVENMSILGNVYVCDIYCNCGVVLNHNKEIKSDILKP 360  
Qy 358 EIVM 361  
Db 361 DIVM 364

RESULT 3  
P85357

GDP-mannose pyrophosphorylase like protein [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cross)  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 02-Mar-2001  
C:Accession: P85357  
R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp  
NATURE 402, 769-777, 1999  
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.  
A:Reference number: A85001; MUID:20083488; PMID:10617198  
A:Accession: P85357  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-351 <STO>  
A:Cross-references: GB:NC\_001268; NID:g7269958; PIDN:CAB79775.1; GSPDB:GN00140  
C:Genetics:  
A:Gene: AT4g30570  
A:Map position: 4  
C:Superfamily: mannose-1-phosphate guanylyltransferase

Query Match 73.8%; Score 1376; DB 2; Length 351;  
Best Local Similarity 76.3%; Pred. No. 1.4e-101;  
Matches 264; Conservative 44; Mismatches 36; Indels 2; Gaps 2;  
Qy 1 MKALILVGGCGTRLRPLTSLSPFKPLVDFAFKPMILHQIEALKEVGYTEVVLAINY-repE 59  
Db 1 LKALILVGGCGTRLRPLTSLSPFKPLVDFAFKPMILHQIEALKEVGYTEVVLAINHQOPEV 60  
Qy 60 MINFLKDFEDKLGITITCSQETPLGTAGPLALARDKLADGSGQFFVNLSDVISEYPFA 119  
Db 61 MLNPFVKEYEKLKIKITFSQETPLGTAGPLALARDKLVDGSGQFFVNLSDVISEYPLL 120  
Qy 120 ELIKFKHCHGGEATIMVTKVDEPSKGVVWMEETGVRERVEKPIFVGKNINAGIYLL 179  
Db 121 EMIEFHKTNRAEASIMVTEVDDPSKYGVVTEGTARVESFEVKPKHFVGNKINAGIYLL 180  
Qy 180 NPSVLDRIELRPTSIKEIEFPQIAADQOLYAMVLPGFWMIDGQPRDYITGLRLYDLSIRK 239  
Db 181 SPVSLDRIELRPTSIKEIEFPQIAAEKLYAMVLPGFWMIDGQPRDYITGLRLYDLSLRE 240  
Qy 240 KSAKLATGAHVGNVNLVHESAKIGECGLIGPDVAIGPGCVVEDGVRLSRTVMRGVRIK 299  
Db 241 KTOELATGNIIGNVNLVHESAVICEGLIGPDVVIIGPGCVVDGVRLEFGCTVMRGVWIK 300  
Qy 300 KHCISNSIIIGWSTVQGWARIENMTILGEDVHVCD-DEVYNSNGGV 344  
Db 301 EHACISNSIVGWDSTVGRWARVENITVLGKDVYVADADEVYNSGVYI 346

RESULT 4  
S67590  
mannose-1-phosphate guanylyltransferase (EC 2.7.7.13) - yeast (Saccharomyces cerevisi  
N:Alternate names: protein D2566; protein YDL055c; PSA1 protein  
C:Species: Saccharomyces cerevisiae  
C:Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 20-Jun-2000  
C:Accession: S67590; S59737; S59374; S62744  
R:Blöcker, H.; Brandt, P.  
submitted to the Protein Sequence Database, July 1996  
A:Reference number: S67587  
A:Accession: S67590  
A:Molecule type: DNA  
A:Residues: 1-361 <BLD>  
A:Cross-references: EMBL:Z74103; NID:g1431052; PIDN:CAA98617.1; PID:g1431053; MIPS:YD  
A:Experimental source: strain S288C  
R:Schultz, J.; Sprague, G.F.  
submitted to the EMBL data Library, April 1995  
A:Description: Isolation of the gene encoding mannose-1-phosphate guanylyltransferase f  
A:Reference number: S59737  
A:Accession: S59737  
A:Molecule type: DNA  
A:Residues: 1-49, 'A', 51-361 <SCH>  
A:Cross-references: EMBL:U24437; NID:g894203; PIDN:AAA69677.1; PID:g894204  
R:Renton, B.K.; Plump, S.D.; Roos, J.; Lennarz, W.J.; Cross, P.R.  
submitted to the EMBL data Library, January 1995  
A:Description: Overexpression of 5. Cerevisiae G1 cyclins restores viability of algi



A:Reference number: S59374

A:Accession: S59374

A:Molecule type: DNA

A:Residues: 1-15,'G',17-361 <BEN>

A:Cross-references: EMBL:U19608

R:Benton, B.K.; Plump, S.D.; Roos, J.; Lennarz, W.J.; Cross, F.R.

Curr. Genet. 29, 106-113, 1996

A:Title: Over-expression of *S. cerevisiae* G(1) cyclins restores the viability of *alg1* N

A:Reference number: S62744; MUID:96418871; PMID:8821656

A:Accession: S62744

A>Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-361 <BEW>

A:Cross-references: EMBL:U19608; NID:g1292897; PIDN:AAC49289.1; PID:g1292898

C:Genetics:

A:Gene: SGD:PSA1; MPG1

A:Cross-references: SGD:S0002213; MIPS:YDL055c

A:Map position: 4L

C:Superfamily: mannose-1-phosphate guanylyltransferase

C:Keywords: nucleotidyltransferase

Query Match 63.2%; Score 1178; DB 2; Length 361;  
Best Local Similarity 61.0%; Pred. No. 7.5e-86;  
Matches 221; Conservative 56; Mismatches 83; Indels 2; Gaps 2;

QY 1 MKALILVGGFGRRLRLTLSPKPLVDFAFKPMILHQIEALKEVGTVTVLAINRPEVM 60

Db 1 MKGILVGGYGRRLRLTLTPKPLVEFGNRPMLHQIEALANAGVTDIVLAVNRPPEVM 60

QY 61 INFLKDFEDKLGITTCQETEPGLTAGLALARDKLDGSGQPFVFLNSDVISEYPAE 120

Db 61 VETLKKYEKGVNITVSVEPIETAGPLKLAEDVLKK-DNSPFFVFLNSDVISEYPAE 119

QY 121 LIKFKHCHGGEATIMVTKVDEPSKYGVVMEET-GRVERFEVKPKIFVGNKINAGIYLL 179

Db 120 LADPHKAGGKGTATVKVDEPSKYGVIVHDIAENPLNIDREVEKPEFVGNRINAGLYL 179

QY 180 NPSVLDRIELRPTSTIEKEVFPQIAADQOLYAMVLPFGFMDVGQPRDYITGLRLYLDSTRK 239

Db 180 NPEVIDIEMKPTSTIEKETFPILVEEKQLYSFDLEGFMDVGQPKDFUSGVTVLYLNSLAK 239

QY 240 KSAAKLATGARVGNLVHESAKIGEGCLIGPDVAIGPCVVEDGVRLSRTVMRGVRIK 299

Db 240 RQPKLATGANIVGNALIDTPAKTSSTAKIGPDVVGPNVTIGDVRITRSVVLNCSITIK 299

QY 300 KHACISNIICWHSTVGQWARIENMTILGEDVHVCDEYVNSGVVLPKHEIKSSILKPEI 359

Db 300 NPSLVKSTIVGNSTVGQWCRLEGVTVLGDDEVEKDEIYINGGVLPKHSISDNVPKEAI 359

QY 360 VM 361

Db 360 IM 361

RESULT 5

T41209

mannose-1-phosphate guanylyltransferase - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 21-Jan-2000

C:Accession: T41209

R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Murphy, L.; Harris, D.

submitted to the EMBL Data Library, September 1998

A:Reference number: Z21969

A:Accession: T41209

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-363 <LYN>

A:Cross-references: EMBL:AL031538; PIDN:CAA20770.1; GSPDB:GN00068; SPDB:SPCC1906.01

A:Experimental source: strain 972h-; cosmid c1906

C:Genetics:

A:Gene: SPDB:SPCC1906.01

A:Map position: 3

C:Superfamily: mannose-1-phosphate guanylyltransferase

Query Match 63.0%; Score 1175; DB 2; Length 363;

Best Local Similarity 62.1%; Pred. No. 1.3e-85;

Matches 226; Conservative 54; Mismatches 80; Indels 4; Gaps 4;

QY 1 MKALILVGGFGRRLRLTLSPKPLVDFAFKPMILHQIEALKEVGTVTVLAINRPEVM 60

Db 1 MKALILVGGFGRRLRLTLTPKPLVEFGNRPMLHQIEALAAAGVTDIVLAVNRPPEIM 60

QY 61 INFLKDFEDKLGITTCQETEPGLTAGLALARDKLDGSGQPFVFLNSDVISEYPAE 120

Db 61 VEALAKYEKEYNVNITVSVENEPLTAGLALARDILAK-DHSPPFFVFLNSDVISEYPAE 119

QY 121 LIKFKHCHGGEATIMVTKVDEPSKYGVV-MEETATGRVERFEVKPKIFVGNKINAGIYLL 179

Db 120 LAAFHKAHGAEGTIVVTKVEEPSKYGVVHYHPNSIESLIERFEVKPEFVSNRINGGIYIL 179

QY 180 NPSVLDRIELRPTSTIEKEVFPQIAADQOLYAMVLPFGFMDVGQPRDYITGLRLYLDSTRK 239

Db 180 NPSVLDRIEPRPTSTIEKEVFPAMVNDKQLHSPFDEGWMVDGQPKDYLTGTCILYLSLRK 239

QY 240 KSAAKLA-TGAHVGNLVHESAKIGEGCLIGPDVAIGPCVVEDGVRLSRTVMRGVRI 298

Db 240 HKPEILAPASSNLIIGNVLIDPSATIGKCKIGPNVVGPNVTIGDGVRLQRCALIKSSRV 299

QY 299 KHACISNIICWHSTVGQWARIENMTILGEDVHVCDEYVNSGVVLPKHEIKSSILKPE- 357

Db 300 RDHAWYKSSIVGNSTVGLSWSRLNSVLDGDDVVVNDIYVNGSILPHKHSISANIEVPG 359

QY 358 EIVM 361

Db 360 TIVM 363

RESULT 6

T42371

probable mannose-1-phosphate guanylyltransferase (EC 2.7.7.13) - fission yeast (Schiz

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 21-Jul-2000

C:Accession: T42371

R:Yoshioka, S.; Kato, K.; Nakai, K.; Okayama, H.; Nojima, H.

DNA Res. 4, 363-369, 1997

A:Title: Identification of open reading frames in Schizosaccharomyces pombe cDNAs.

A:Reference number: Z17323; MUID:98162722; PMID:9501991

A:Accession: T42371

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-363 <YOS>

A:Cross-references: EMBL:D89128; NID:g1749463; PIDN:BAAL3790.1; PID:g1749464

A:Experimental source: strain PR745

C:Function:

A:Description: generates GDP-mannose and pyrophosphate from mannose-1-phosphate and G

C:Superfamily: mannose-1-phosphate guanylyltransferase

C:Keywords: nucleotidyltransferase

Query Match 62.4%; Score 1164; DB 2; Length 363;

Best Local Similarity 61.5%; Pred. No. 9.7e-85;

Matches 224; Conservative 55; Mismatches 81; Indels 4; Gaps 4;

QY 1 MKALILVGGFGRRLRLTLSPKPLVDFAFKPMILHQIEALKEVGTVTVLAINRPEVM 60

Db 1 MKALILVGGFGRRLRLTLTPKPLVEFGNRPMLHQIEALAAAGVTDIVLAVNRPPEIM 60

QY 61 INFLKDFEDKLGITTCQETEPGLTAGLALARDKLDGSGQPFVFLNSDVISEYPAE 120

Db 61 VEALAKYEKEYNVNITVSVENEPLTAGLALARDILAK-DHSPPFFVFLNSDVISEYPAE 119

QY 121 LIKFKHCHGGEATIMVTKVDEPSKYGVV-MEETATGRVERFEVKPKIFVGNKINAGIYLL 179

Db 120 LAAFHKAHGAEGTIVVTKVEEPSKYGVVHYHPNSIESLIERFEVKPEFVSNRINGVIYIL 179

QY 180 NPSVLDRIELRPTSTIEKEVFPQIAADQOLYAMVLPFGFMDVGQPRDYITGLRLYLDSTRK 239

Db 180 NPSVLDRIEPRPTSIEKEVFPAMVNDKQLHSFDLEGYMDVQPKDYLTGCTCLYLSLRK 239  
QY 240 KSAAKLA-TGAHVGVNVLVHESAKIGEGCLIGPDVAIGPCVVEDGVRLSRCTVMRGVRI 298  
Db 240 HPEILAPASSNIGNVLDIPATIKCKIGENYVYIGNYIGDGYRQRCALIKSSRV 239  
QY 299 KKHACISNIHGHSTVGQWARIENMTILGEDVHVCDVYVSGVVLPHKEIKSSILKP- 357  
Db 300 RDHAWKSSIVGNWSTLGSWSRLNVSVLGDDVVVNDVYVNGSGSLPHKKSISANIEVPG 359  
QY 358 EIVM 361  
Db 360 TIVM 363  
RESULT 7  
T32900  
hypothetical protein C42C1.5 - *Caenorhabditis elegans*  
C/Species: *Caenorhabditis elegans*  
C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 21-Jan-2000  
C/Accession: T32900  
B/Murray, J.; Rohlfing, M.; Davidson, S.  
submitted to the EMBL Data Library, January 1998  
A/Description: The sequence of *C. elegans* cosmid C42C1.  
A/Reference number: Z1243  
A/Accession: T32900  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-373 <MUR>  
A/Cross-references: EMBL:AF043695; PIDN:AB97947.1; GSPDB:GN00019; CESP:C42C1.5  
A/Experimental source: strain Bristol N2; clone C42C1  
C/Genetics:  
A/Map position: 1  
A/Introns: 1/3; 269/1; 308/3; 343/3  
C/Superfamily: mannose-1-phosphate guanylyltransferase  
Query Match 56.1%; Score 1045; DB 2; Length 373;  
Best Local Similarity 55.3%; Pred. NO. 2.8e-75;  
Matches 203; Conservative 64; Mismatches 92; Indels 8; Gaps 3;  
QY 1 MKALILVGGFGTNRPLTSLFPKPLVDFAFKPMILHQLKEKVGTEVVLAINYRPEVM 60  
Db 9 MKALILVGGYGTNRPLTTLTPKPLVEFAFKPMMLHQLKEALAEVGDVTVLAVSYRAEQ 68  
QY 61 INFLKDFEDKLGITITCSQETPLTAGPLALARDKADSGOPFFVLNSDVISEYPPAE 120  
Db 69 EQEMTVHADRLGVKLIFSLSEEPPLGTAGPLALARKHL-EGDA-PFFVLNSDVICDFFPK 126  
QY 121 LKFKHCKHGEATIMTVKDEPSKVGVMVEEATGRVERFVKPFVGNKINAGVYLLN 180  
Db 127 MYEFHNKHEGTTIATVKVEPSKYGVVVFDQDKGIDDFEVEKQYVGNKINAGIYIFS 186  
QY 181 PSVLDRIELRPTSIEKEVFPQIAADQQLYAMVLPFGFMDVQGPDRDITGLRLYLDLSIRK 240  
Db 187 SKILDRIPKPTSIEKEIFPEMAFNSNLVAFVLPFGFMDVQGPDKDFLKNHCHTT 246  
QY 241 SAAKATG-----AHVGVNVLVHESAKIGEGCLIGPDVAIGPCVVEDGVRLSRCTVMR 294  
Db 247 KSKDLTGTSNIHTATIRGNVYDPATYGENCVIGFDVYVYIGRVKIEGGYRLHSTILS 306  
QY 295 GVRIKHKACISNIHGHSTVGQWARIENMTILGEDVHVCDVYVSGVVLPHKEIKSSI 354  
Db 307 DSSIGNYSWWSGVSGKKGCHIGSWKRIENICVIGDDVVVKDELILNGASVLPKHSIAVNV 366  
QY 355 LRPEIVM 361  
Db 367 PSKDIIM 373  
RESULT 8  
G96778  
hypothetical protein F9E10.24 [imported] - *Arabidopsis thaliana*

C/Species: *Arabidopsis thaliana* (mouse-ear cress)  
C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C/Accession: G96778  
R/Rheologis, A.; Ecker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alon  
ansen, N.B.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,  
Nature 408, 816-820, 2000  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marzla  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo  
kei, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Daylis, R.W.  
A/Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
A/Reference number: A86141; MUID:21016719; PMID:11130712  
A/Accession: G96778  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-415 <SYO>  
A/Cross-references: GB:AE005173; NID:G6646773; PIDN:AAF21085.1; GSPDB:GN00141  
C/Genetics:  
A/Gene: F9E10.24  
A/Map position: 1  
Query Match 29.5%; Score 550; DB 2; Length 415;  
Best Local Similarity 31.0%; Pred. NO. 6.1e-36;  
Matches 136; Conservative 86; Mismatches 146; Indels 48; Gaps 9;  
QY 3 ALIIVGG--FGTNRPLTSLFPKPLVDFAFKPMILHQLKEKVGTEVVLAINYRPEV 59  
Db 11 AVIMVGGTGTNRFPGLNLPKPLFPIAGQPPVHHFISACKRPNLQAQIYLVGYEERE 70  
QY 60 MINFLKDFEDKLGITITCSQETPLTAGPLALARDKADSGOPFFVLNSDVISEYPPAE 119  
Db 71 FALVVSATSNELKVPVRYLRDKPHGSAGGLYHFNLMEDSPSHIFLLNCDCSFPPLP 130  
QY 120 ELIKFKHCKHGEATIMTVK--DEPSKYGVVYMEEATGRVERFVKPFVGNKINAGIY 177  
Db 131 KMLEAHRGYGGTGLIVIKVSPESASQFGLVADPVTNELLHVTKEPTEFVSDRINCQV 190  
QY 178 LLNPSVL-----DRIELR-----PTS---IEKEVFPQIAADQQLY 209  
Db 191 VTFTEIFNAGDVSTQRKDRATLAKVSFEALQAPTRITDFVRLDQDILSPLAGKRLY 250  
QY 210 AMVLPFGFMDVQGPDRDITGLRLYLDLSIRKSKAAKATG-----AHVGVNVLVHESAKIG 264  
Db 251 TYETMDWEQIKSPGMSLRCSGLYLSQFRLTSPOLLASGDGTRSAIVIGDYIHPSAKVH 310  
QY 265 EGCILIGPDVAIGPCVVEDGVRLSRCTVMRGVRIKHKACISNIHGHSTVGQWARI-- 322  
Db 311 PTAKIGPNVSTANARVCGVRLMSCIILDVETMENAVVTNAIVGWKSSITGRMSRVOAE 370  
QY 323 -----NMVILGEDVHVCDVYVSGVVLPHKEIKSSILKPEIVM 361  
Db 371 GYVNSKLGVTILGDSVAVEDEVVYVTSIVLPNKTILNVSV-QDEIIL 415  
RESULT 9  
AB2101  
mannose-1-phosphate guanylyltransferase [imported] - *Nostoc* sp. (strain PCC 7120)  
C/Species: *Nostoc* sp.  
A/Note: *NOSTOC* SP. Strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120  
C/Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002  
C/Accession: AB2101  
R/Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriqu  
Nakasaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata  
DNA Res. 9, 205-213, 2001  
A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing *Cyanobacterium*  
A/Reference number: AB1807; MUID:21595285; PMID:11759840  
A/Accession: AB2101  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-842 <KUR>  
A/Cross-references: GB:BA000019; PIDN:BA074060.1; PID:gl7131453; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: alr2361

Query Match 27.0%; Score 504; DB 2; Length 842;  
Best Local Similarity 32.6%; Pred. No. 7.1e-32;  
Matches 129; Conservative 83; Mismatches 116; Indels 68; Gaps 16;

QY 1 MKALILVGGFGRTRPLTLSPKPLVDFAFKPMILHQIETALKEVGTETWVLAINRPEVM 60  
DB 1 MRVLMAGGSGTRLRPLTCDLPKWPVILNRPPIAEHIINLLKRHHITEVIATLHLPDVL 60  
QY 61 INFLKDFEDKLGITITCS-QETPLGTAGPLALARDKLADGGQPFVLSNDSVISEYPPFA 119  
DB 61 RYFQDGGD-FGVQMTVAVEDQPLGTAGCV----KNIAELLDEFFLVISGDSITDFDLG 115  
QY 120 ELIFKHCHGGEATIMVTKVDEPSKYGVVVMEEATGRVERFEKRP----KIFVGNKINAGI 176  
DB 116 EAIAPFKQKQSKATLILTRVNPPIEFVGVITDEA-GKIKRFLERKSTSEIF-SDTVNRTG 173  
QY 177 YLLNPSVLDRLTELPSTSE----KEVFP-QIAADQOLYAMVLPGFMDVQGPDRDY----I 227  
DB 174 YILEPEVL----EYLPSTNECDFSKDLFPLLAKDEPMPTGYVAEGVCDVGHLDAYREAOY 230  
QY 228 TGL-RLYLDGIRKKSAAKATG------AHVVGNYLVHESAKIG------EGCLIGP 271  
DB 231 DGLDRKVKLDFARENSPGLWIGQNTYIDPSAHTAPAVIGNNCRIGARVQIEAGTVIGD 290  
QY 272 DVAITGP------GCVVEDGVRLSCTVMRGVRIKKHIA-CISNSIIGHSTVGOWA 319  
DB 291 NVITGADANLRPIVWNGAIGEEAQLSACVISRGRVDRRAHVLEASVWGSLSVGBEA 350  
QY 320 RIENWTLIGEDVHVCDVEYVSGVGVLPHPKIKSSIL 355  
DB 351 QI------SPGVRWPSKKIESGAI 369

RESULT 10

G84459  
probable GDP-mannose pyrophosphorylase [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L.;  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487; PIDN:AAD22341.1; GSPDB:GN00139  
A:Accession: G84459  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-385 <STO>  
A:Cross-references: GB:AE002093; NID:g4544432; PIDN:AAD22341.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: At2g04650  
A:Map position: 2

Query Match 26.5%; Score 493.5; DB 2; Length 385;  
Best Local Similarity 30.2%; Pred. No. 1.7e-31;  
Matches 118; Conservative 80; Mismatches 148; Indels 45; Gaps 8;

QY 3 ALILVGG-FGTRLRPLTLSPKPLVDFAFKPMILHQIETALKEV-GVTEVVLAINRPEV 59  
DB 8 AVIMVGPTKTRFRPLSFNTPKPLIPLAGPMIHHPIPSACKKISNLAQIIFLIGYEER 67  
QY 60 MINFLKDFEDKLGITITCSQTEPLGTAGPLALARDKLADGGQPFVLSNDSVISEYPPFA 119  
DB 68 FALVSSISNELKIPVRYLKEDKPHGSGAGALYFDRIMEEKPSHVFLNCDVCCSFLLQ 127  
QY 120 ELIFKHCHGGEATIMVTKV-DEPSKYGVVVMEEATGRVERFEKRPKIFVGNKINAGIY 177  
DB 128 GILDAHRRYGGIGTWMVTKVSAEASQFGELIADPDTKELLHYTEKPEFTVSDLINCGY 187

QY 178 LLNPSVLDRIE------LRPTS------IEKEVFPQIAADQOLYAMVLPGF 216  
DB 188 VFTSDIFNAIEEYVSQIRDSSNYQSATRSVPADFVRLDQDILSPLAGKQLYTYENKDF 247  
QY 217 WMDVGOPRDIYIGRLRYLDSIRKKSAAKATG------AHVVGNYLVHESAKIGEGCLIG 270  
DB 248 WEQIKTPGKSLCALSALYSQFRETSPHILASGDGTNRKPTIIGDVYIHPVKLHPTAKIG 307  
QY 271 PDVAITGCGVVEDGVRLSCTVMRGVRIKKHACISINSIIGHSTVGOWARIENWTLIGED 330  
DB 308 PNVSISANRVGSGVRLISCIILDDVEIKENAVVINSIIGKSSIGEAVTVE------ 359  
QY 331 VHVCDVEYVSGVGVLPHPKIKSSILKPEIVM 361  
DB 360 ----DEVAVIGSIVLQNKTLNVSV-QDDIIL 385  
RESULT 11  
A70363  
mannose-1-phosphate guanylyltransferase - Aquifex aeolicus  
C:Species: Aquifex aeolicus  
C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 05-Nov-1999  
C:Accession: A70363  
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;  
V.  
Nature 392, 353-358, 1998  
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
A:Reference number: A70300; MUID:98196666; PMID:9537320  
A:Accession: A70363  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-831 <AQF>  
A:Cross-references: GB:AE000704; NID:g2983301; PIDN:AAC06893.1; PID:g2983302; GB:AE00  
A:Experimental source: strain VF5  
C:Genetics:  
A:Gene: mpg

Query Match 26.4%; Score 492; DB 2; Length 831;  
Best Local Similarity 31.2%; Pred. No. 6.2e-31;  
Matches 123; Conservative 86; Mismatches 133; Indels 52; Gaps 14;  
QY 1 MKALILVGGFGRTRPLTLSPKPLVDFAFKPMILHQIETALKEVGTETWVLAINRPEVM 60  
DB 1 MKGVILAGGFGTRIQPLTNSIPKMLPVANRPIMEHVHRLKEAGIEEIVLLYQAEVI 60  
QY 61 INFLKDFEDKLGITITCSQTEPLGTAGPLALARDKLADGGQPFVLSNDSVISEYPPFAE 120  
DB 61 KNVFKDGS-D-FGVKIIVYQPEADYGTAGAVKQAOYVL----NETFIVSGVITDFNLSE 115  
QY 121 LIKFHKCHGGEATIMVTKVDEPSKYGVVVMEEATGRVERFEKRP----KIFVGNKINAGIY 177  
DB 116 LIAFHKSCKSKFTLALYSVENPLQFGVVIITNK-EGKVLKLEKPGWGEVF-SDTVNRTG 173  
QY 178 LLNPSVLDRI-ELRPTSIKEVFPQI-AADQOLYAMVLPGFMDVQ------PRDIYT 228  
DB 174 VVEPEILNYPEDKPFDFAMDLFPKLMKSGIDLWALKMRGYWRDIGNIDSYRVDVHKDIFA 233  
QY 229 GL------RLYLDGIRKKSAAKATGATGAHVVGNYLVHESAKIGEG------CL 268  
DB 234 GLVKTRIPRIRITTKEARIVE------EGTEIPENVSLKGTIVLGNKVKVGESELNKV 288  
QY 269 IGPDAVIGCGVVEDGVRLSCTVMRGVRIKKHACISINSIIGHSTVGOWARIENWTLIG 328  
DB 289 IGNNTVIG------RNKVLFDVSLWNVSIDESEIRNGVICNDVKIGKRVKAKEGVVIA 342  
QY 329 EDVHVCDVEYVSGVGVLPHPKIKSSILKPEIV 360  
DB 343 EDCEVEDEVLFKLDVVVWPEKVIKESGVVTKNIV 376  
RESULT 12  
T39403

probable mannose-1-phosphate guanylyl transferase - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 18-Feb-2000  
C:Accession: T39403  
R:Lyne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.; Badcock, K.; Churcher, C.M.  
submitted to the EMBL Data Library, April 1998  
A:Reference number: 221852  
A:Accession: T39403  
A:Status: preliminary; translated from GB/EMBL/DDRJ  
A:Molecule type: DNA  
A:Residues: 1-414 <L>N>  
A:Cross-references: EMBL:AL022600; PIDN:CAAL8655.1; GSPDB:GN00067; SPDB:SPBC13G1.02  
A:Experimental source: strain 972h; cosmid c13G1  
C:Genetics:  
A:Gene: SPDB:SPBC13G1.02  
A:Map position: 2  
A:Introns: 14/2; 27/3  
C:Superfamily: mannose-1-phosphate guanylyl transferase

Query Match 26.3%; Score 490.5; DB 2; Length 414;  
Best Local Similarity 30.7%; Pred. No. 3.2e-31;  
Matches 127; Conservative 71; Mismatches 157; Indels 59; Gaps 10;

QY 3 ALILVGG--FGRRLRLTLSPFKPLVDFAFKPMILHQAIEALKEVGVTEVVLAINRPEVIM 60  
DB 5 AVILVGGSGRGRFRPLSDVDPKPLKIGREMYYHILAALSKIESYKDFLVGFYDE-- 62

QY 61 INFLDKPEKIGI-----TITCSOTEDIGTACPLALADKLADGSCOPFFVLNSDVIS 114  
DB 63 -SYFKDFINEASHFSPFNRIKYLREYCNLTGTTGGGLYHFRDQLKHTSNVFMHADVCC 121

QY 115 EYFFAEILRFHKGCHGEATIMTKV--DEPSKYGVVVMEEATGRVERFEKPKIFVGNKI 172  
DB 122 SPFOELLNVHHEKKAIVLMATKVKSKEDASNFGCLVEEPTGRVLHVVDKPSYLSNII 181

QY 173 NAGIYLLNFSVLDRI-----ELRP-----TSIEKEVFPQIAAD--QOOLY 209  
DB 182 SCGIYTFDASIFDEIKKAVERRLEEVEKQLRSDEGMEDYLSLTDVLAPLCSDDSKAIY 241

QY 210 AMVLPGFMDVGDOPDYITGLRLYLD-SIRKSKAAKATCAHVGVNVLVHESAKIGGCL 268  
DB 242 AYNTPEFGRQITAGSAVFANSYLQKAYHDGTLKPDTEAEIQQPVFIHFNIAIVSGAK 301

QY 269 IGPDAVAGCCVVEDGVRLSRCTVMRGVRIKHKACISNTHSTVGVQWARIEN----- 323  
DB 302 IGPNSIGARVRIEDGARINSLIOEDCEISANAVLHSLSRCKICKWSRVEGSPILP 361

QY 324 -----MTLGEDVHVCDEYISNGVGVVPLHPKELKSSILKPEIYM 361  
DB 362 SQHSTTRMNSRVQAITVAGACTIVDEVRVQNCVLVLPKELKVGILV-GEIVM 414

RESULT 13  
A70978  
C:Species: Mycobacterium tuberculosis (strain H37Rv)  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
C:Accession: A70978  
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; MUID:98295987; PMID:9634230  
A:Accession: A70978  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-359 <C>L>  
A:Cross-references: GB:Z99771; GB:AL123456; NID:G3242259; PIDN:CAB07091.1; PID:G1877319  
A:Experimental source: strain H37Rv  
C:Genetics:  
A:Gene: rmlA2

C:Superfamily: mannose-1-phosphate guanylyl transferase

Query Match 26.2%; Score 488; DB 2; Length 359;  
Best Local Similarity 34.6%; Pred. No. 4.2e-31;  
Matches 125; Conservative 63; Mismatches 151; Indels 22; Gaps 10;

QY 3 ALILVGGFGRRLRLTLSPFKPLVDFAFKPMILHQAIEALKEVGVTEVVLAINRPEVIM 62  
DB 8 AVILVGGKGRRLRLTLSPAKPMILPTAGLPFLTHLSRIAAGIEHVLIGTSYKPAV--- 64

QY 63 FLKDFED--KLGITITCSQETEPGLTAGPLALARDKADGGQPFVYLNDSVISEYPAE 120  
DB 65 FEAEFGGSGALGLGIEVTEBEHPLOTGGGIANVAKLNDTA--WVFGDVLGADLAQ 121

QY 121 LIKFKHKGCHGEATIMTKVDEPSKYGVVVMEEATGRVERFEKPKIFVGNKINAGIYLLN 180  
DB 122 LLDLHRSNRADYTLQLVRYGDPRAFGCVPTDE-EDRVVAFLEKTEDPTDINAGCYFE 180

QY 181 PSVLDRI-ELRPTSTEKEVFPQIAADQ--QLYAMVLGCFWMDVQOPRDYITGLRLYLD 237  
DB 181 RNVIDRIPQGRVSVREVEFPDALLADGCKIYGVYVDASVWRDMCTPEDFVRG---SADLV 237

QY 238 RKSKAAKATGAHVGVNVLVHESAKIGGCLIGPDVALGPGCVYDGVRLSRCTVMRGVR 297  
DB 238 RGTAPSPALRGR--GEQLVHDGAASPCALLIGTVVVGRAETGPGTRLDGAVIFDGRV 295

QY 298 IKKHACISNTHSTVGVQWARIENMTI-LGEDVHVCDEYISNG-----GVVLPHKEIKS 352  
DB 296 VEAGCVIERSIIGFARLIGPRALLRDGVIGDADIGARCELLSGARWPGVFLPDGGIRY 355

QY 353 S 353  
DB 356 S 356

RESULT 14  
H75070  
C:Species: Pyrococcus abyssi  
C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000  
C:Accession: H75070  
R:anonymous, Genoscope  
submitted to the EMBL Data Library, July 1999  
A:Description: Pyrococcus abyssi genome sequence: Insights into archaeal chromosomes  
A:Reference number: A75001  
A:Accession: H75070  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-413 <KAW>  
A:Cross-references: GB:AJ248286; GB:AL096836; NID:G5458366; PIDN:CAB49069.1; PID:G545  
A:Experimental source: Strain Orsay  
C:Genetics:  
A:Gene: PAB0645  
C:Superfamily: glucose-1-phosphate adenyl transferase

Query Match 25.6%; Score 478; DB 2; Length 413;  
Best Local Similarity 30.7%; Pred. No. 3.2e-30;  
Matches 111; Conservative 87; Mismatches 148; Indels 16; Gaps 6;

QY 1 MKALILVGGFGFGRRLRLTLSPFKPLVDFAFKPMILHQAIEALKEV-GVTEVVLAINRPEV 59  
DB 1 MKAVILAGGFGFGRRLRLTLSPISRRPKPMVPLCKPNLYLLEALEKVEIDEVLSVHYMRGE 60

QY 60 MINPLKDFEDKLGITITCSQETEPGLTAGPLALARDKADGSCOPFFVYLNDSVISEYPAE 119  
DB 61 IREFIQENMRDIPKDIRFYNDPMPLETTGGALKNVEEYSD----DFLYIYDVTNFDFYS 116

QY 120 ELIKFKHKGCHGEATIMTKVDEPSKYGVVVMEEATGRVERFEKPKIFVGNKINAGIYLL 179  
DB 117 ELIEAHKKNGLVVAUTKVVDPDFRFGVITDE-EGKIVEFEKPKPKTKNLVDAGIYMW 175

QY 180 NPSVLDRI-ELRPTSTIEKEVFPQIAADQQLYAMVLPG--FMDVQGRDYITGLRLYLD 236



GenCore version 5.1.3

Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 04:07:11 ; Search time 25 Seconds

(without alignments)

598.918 Million cell updates/sec

Title: US-09-374-967-2

Perfect score: 1864

Sequence: 1 MKALILVGGFGRRLRLTSL.....GVVLPKHKSSILKPEIVM 361

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1178	63.2	361	1	MPGL_YEAST
2	339	18.2	355	1	STRD_STRGR
3	287	15.4	257	1	RFBF_SALTY
4	242	13.0	380	1	GLGC_BACSU
5	233.5	12.5	452	1	E2BG_RAT
6	229.5	12.3	452	1	E2BG_HUMAN
7	228	12.2	246	1	SPSI_BACSU
8	227	12.2	456	1	GLMU_ECOLI
9	224.5	12.0	712	1	E2BE_YEAST
10	223.5	12.0	387	1	GLGC_BACST
11	210	11.3	286	1	RFBA_RHISN
12	209	11.2	283	1	YD34_METJA
13	209	11.2	429	1	GLGC_ANASP
14	209	11.2	456	1	GLMU_HAEIN
15	206.5	11.1	459	1	GLMU_BUCAI
16	204	10.9	439	1	GLGC_SYNY3
17	199.5	10.7	516	1	GLGI_MAIZE
18	198	10.6	272	1	YTDA_BACSU
19	197	10.6	431	1	GLGC_SALTY
20	195.5	10.5	293	1	RFPH_ECOLI
21	190	10.2	430	1	GLGC_ECOLI
22	189.5	10.2	297	1	GALE_ECOLI
23	189.5	10.2	297	1	GALE_SALTY
24	189.5	10.2	298	1	GALE_KLEPN
25	184	9.9	289	1	RLMA_STRMU
26	183.5	9.8	250	1	GLGC_BACCL
27	183	9.8	141	1	YSOD_SULAC
28	182.5	9.8	297	1	GALE_SHIFL
29	181.5	9.7	289	1	RFBA_YEREN
30	180.5	9.7	292	1	RFBA_SALAN
31	180.5	9.7	456	1	GCAD_BACSU
32	178	9.5	288	1	RFBA_NEIMB
33	177	9.5	293	1	RBA2_ECOLI

34	176.5	9.5	279	1	GAUL_PSEAE	Q91291 pseudomonas
35	176.5	9.5	292	1	RFBA_SALTY	P26393 salmonella
36	176	9.4	288	1	RFBA_NEIMA	P37762 neisseria g
37	176	9.4	288	1	RFBA_NEIMA	P57040 neisseria m
38	175.5	9.4	295	1	RFBA_XANCP	P55256 xanthomonas
39	172.5	9.3	292	1	GTAB_BACSU	Q05852 b utp--gluc
40	171.5	9.2	404	1	E2BG_CAEEL	P80361 caenorhabdi
41	171	9.2	483	1	GLGS_ORYSA	P15280 oryza sativ
42	170.5	9.1	293	1	RBA1_ECOLI	P37744 escherichia
43	169.5	9.1	291	1	GAUL_MYCPN	P75124 mycoplasma
44	169.5	9.1	473	1	GLGS_WHEAT	P30523 triticum ae
45	169.5	9.1	513	1	GLGS_HORVU	P55238 hordeum vul

## ALIGNMENTS

RESULT 1

ID	MPGL_YEAST	STANDARD;	PRT;	361 AA.
AC	P41940;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Mannose-1-phosphate guanylttransferase (EC 2.7.7.13) (ATP-mannose-1-phosphate guanylttransferase) (NDP-hexose pyrophosphorylase).			
GN	MPGL OR PSA1 OR YDL055C.			
OS	Saccharomyces cerevisiae (Baker's yeast).			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.			
OX	NCBI_TaxID=4932;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96418871; PubMed=8821656;			
RA	Benton B.K., Plump S.D., Roos J., Lennarz W.J., Cross F.R.;			
RT	"Over-expression of S. cerevisiae G1 cyclins restores the viability of alg1 N-glycosylation mutants.";			
RL	Curr. Genet. 29:106-113(1996).			
[2]				
RP	SEQUENCE FROM N.A.			
RA	Schultz J., Sprague G.F. Jr.;			
RL	Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.			
[3]				
RP	SEQUENCE FROM N.A.			
RA	Bloecker H., Brandt P.;			
RL	Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.			
CC	-!- FUNCTION: ESSENTIAL PROTEIN WHICH MAY BE INVOLVED IN THE REGULATION OF CELL CYCLE PROGRESSION.			
CC	-!- CATALYTIC ACTIVITY: GTP + alpha-D-mannose 1-phosphate -> diphosphate + GDP-mannose.			
CC	-!- SIMILARITY: TO S-TYPHIMURIUM CDP-GLUCOSE PYROPHOSPHORYLASE (RFBF).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; U19608; AAC49289.1; -			
DR	EMBL; U24437; AAA69677.1; -			
DR	EMBL; Z74103; CAA98617.1; -			
DR	SGD; S0002213; PSA1.			
DR	InterPro; IPR001825; NTP_transferase.			
DR	Pfam; PF00132; hexapep. 4.			
DR	Pfam; PF00483; NTP_transferase; 1.			
KW	Transferase; Kinase; Cell cycle.			
FT	CONFLICT 50 50 V -> A (IN REF. 2).			
SQ	SEQUENCE 361 AA; 39566 MW; DBFIC39BEAE0B776 CRC64;			

Query Match 63.2%; Score 1178; DB 1; Length 361;  
Best Local Similarity 61.0%; Pred. No. 1.7e-85;

[illegible]



```
CC BIOSYNTHESIS PATHWAY OF LIPOPOLYSACCHARIDE BIOSYNTHESIS.
CC -!- SIMILARITY: TO MANNOSYL-1-PHOSPHATE GUANYLTRANSFERASE (MPG1).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X56793; CAA40120.1; -.
DR EMBL; AE008792; AAL20996.1; -.
DR PIR; S15304; S15304.
DR StyGene; SG10346; rfbF.
DR InterPro; IPR001825; NTP_transferase.
DR Pfam; PF00483; NTP_transferase; 1.
KW Lipopolysaccharide biosynthesis; Transferase; Kinase;
KW Complete proteome.
SQ SEQUENCE 257 AA; 29035 MW; 04DB33D51357D70E CRC64;

Query Match 15.4%; Score 287; DB 1; Length 257;
Best Local Similarity 28.5%; Pred. No. 1.4e-15;
Matches 76; Conservative 50; Mismatches 109; Indels 32; Gaps 7;

QY 1 MKALILVGGTGLRLPLTLSPKPLVDFANKPMILHIOEALKEGVGVTEVVLAINYRPEVM 60
DB 1 MKAVILAGGLTGLRLEETIYKPKPMVEIGGKPLIWHIMKMYSVHGINKDFIICCGYKGVY 60
QY 61 INFKDP-----EDKLG-----TITCQETPLGTAGPLALARDKLDG 100
DB 61 KEYFANYFLHMSDVTFHMAENRMEVHKRVEPNWTLVDGDSMTGRLKRAEYVKD- 119
QY 101 SGQFFVLNSDVISEYPPFAELIKFKHCHGGEATIMVTKVDPSKYGVVMEATGRVERF 160
DB 120 -DEAFLFTYGDGVADLIDKATIDFKAHGKKATL--TATFPGRFG--ALDIRAGQVRSE 174
QY 161 VERPKIFVGNKINAGIYLLNPSVLDRELPTSTSEKEVFPQIADQQLYANVLPGFMDV 220
DB 175 QEAPK-GDGAMINGGVFLNPSVIDLIDNATWEQPLMTLAQQGELMAFEHFGFQPM 233
QY 221 GPRDYITGLRLYLDIRKSKAAKAT 247
DB 234 DTLRD-----KVLEGLWEKGPAPWT 255

RESULT 4
GLGC_BACSU STANDARD; PRT; 380 AA.
AC GLGC_BACSU STANDARD; PRT; 380 AA.
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Glucose-1-phosphate adenyltransferase (EC 2.7.7.27) (ADP-glucose
DE synthase) (ADP-glucose pyrophosphorylase).
GN GlGC.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=94195107; PubMed=8145641;
RA Kiel J.A.K.W., Boels J.M., Beidman G., Venema G.;
RT "Glucose-1-phosphate adenyltransferase: molecular characterization of an
RT operon encoding enzymes involved in glycogen biosynthesis and
RT degradation.";
RL Mol. Microbiol. 11:203-218(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98048467; PubMed=9387221;
RA Lapidus A., Galleron N., Sorokin A., Ehrlich S.D.;
```

```
RT "Sequencing and functional annotation of the Bacillus subtilis genes
in the 200 kb rnb-dnaB region.";
RL Microbiology 143:3431-3441(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertsch M.G., Bessieres P., Bolotin A., Borchert S.,
RA Brouillet S., Bruschi C., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghm S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guiseppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinols S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivalta C., Rocha E., Roche B., Rose M., Sadate Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serro P., Shin B.S., Soldo B.,
RA Sorokin A., Taccini E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
subtilis";
RL Nature 390:249-256(1997).
CC -!- CATALYTIC ACTIVITY: ATP + alpha-D-glucose 1-phosphate =
CC diphosphate + ADP-glucose.
CC -!- PATHWAY: Glycogen biosynthesis; first step.
CC -!- INDUCTION: EXPRESSED EXCLUSIVELY ON MEDIA CONTAINING CARBON
CC SOURCES THAT ALLOW EFFICIENT SPOULATION.
CC -!- SIMILARITY: BELONGS TO THE BACTERIAL AND PLANTS GLUCOSE-1-
CC PHOSPHATE ADENYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z25795; CAA81041.1; -.
DR EMBL; AF008220; AAC00215.1; -.
DR EMBL; Z99119; CAB15075.1; -.
DR PIR; S36625; S36625.
DR Subtilist; BG10908; glgc.
DR InterPro; IPR001825; NTP_transferase.
DR Pfam; PF00483; NTP_transferase; 1.
DR PROSITE; PS00808; ADP_GLC_PYROPHOSPH_1; 1.
DR PROSITE; PS00809; ADP_GLC_PYROPHOSPH_2; 1.
DR PROSITE; PS00810; ADP_GLC_PYROPHOSPH_3; 1.
KW Glycogen biosynthesis; Transferase; Nucleotidyltransferase;
KW Complete proteome.
SQ SEQUENCE 380 AA; 42602 MW; 8383169C727A4DB5 CRC64;

Query Match 13.0%; Score 242; DB 1; Length 380;
Best Local Similarity 24.3%; Pred. No. 8e-12;
Matches 92; Conservative 74; Mismatches 178; Indels 34; Gaps 13;

QY 3 ALILVGGTGLRLPLTLSPKPLVDFANKPMIL-HQIEALKEGVGVTEVVLAINYRPEVM 61
```



```

Db 7  AMLIAGGKSRUSGUTKWKNAKFAVGGKRYRIDFTLSNCSNSGIDTVGILLTIQPLELN 66
Qy 62  NPLK-----DFED-KLGTIT-----TCSQETEPFLATGAPLALARDKLADSCGPPFVLNS 110
Db 67  SVIGIGSAWLDYNGCVTVLPYAESSEVKWKGTASSIEMNLVNOYDPEYVILISG 126
Qy 111 DVISYPAELIKFKHGGEATIMYTKV--DEPSKYGVVYMEATGRVERFEKPKFY 168
Db 127 DHIVMDYCKMLDYTERKADVTISVIEGWEEASRFQ-IMKANDPDGTTTHFDKPKPK 185
Qy 169 GNKNAGIYLLN-PSVLDRIEL---RPTS---IEKEVFQIARAD-QOLYAMVLPFGFMDY 220
Db 186 SMLASHGRIIFNWLKQVLEDDQNPYSYSHDFGRDIIIFLLLEKRLKLSAYPFKGYWKDV 245
Qy 221 GOPRDYITGLRLYLSIRKSAAKATGCAHVVGCVNLVSHESAKITGECGLICPDVAIGPGCV 280
Db 246 GTVQSL---WEANMDLLKEDSELKUFERKWKIYSVNPQPPQFISSDAQVDSLVNEGCV 302
Qy 281 VEDGVLSRCTVNRGVRIRKKHACISNIIGWSTVGQWARIENMTILGEBDVHVCDEVYSN 340
Db 303 VYGNV--SHSVLFQGVTVGKHTVTSSVIMPDPVTIGEHVVVIAIYVN-----GMVLDP 354
Qy 341 GGVLPKHEIKKSILKPE 358
Db 355 GAVIRSEKXDEIEVLLVSE 372

RESULT 5
E2BG_RAT STANDARD: PRT: 452 AA.
AC E2BG_RAT
ID E2BG_HUMAN STANDARD: PRT: 452 AA.
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Translation initiation factor eIF-2B gamma subunit (eIF-2B GDP-GTP
DE exchange factor).
DE EIF2B3 OR EIF2BG.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Skeletal muscle;
RX MEDLINE=96404916; PubMed=8809057;
RA Price N.T., Kimball S.R., Jefferson L.S., Proud C.G.;
RT Cloning of cDNA for the gamma-subunit of mammalian translation
RT initiation factor 2B, the guanine nucleotide-exchange factor for
RT eukaryotic initiation factor 2.;
RL Biochem. J. 318:631-636(1996).
CC -!- FUNCTION: CATALYZES THE EXCHANGE OF EUKARYOTIC INITIATION FACTOR
CC -!- 2-BOUND GDP FOR GTP.
CC -!- SUBUNIT: COMPLEX OF FIVE DIFFERENT SUBUNITS; ALPHA, BETA, GAMMA,
CC DELTA AND EPSILON.
CC -!- TISSUE SPECIFICITY: BRAIN, HEART, LIVER, MUSCLE AND TESTES.
CC -!- SIMILARITY: BELONGS TO THE EIF-2B GAMMA/EPSILON SUBUNITS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U38253; AAC52788.1; .
DR InterPro: IPR001825; NTP transferase.
DR pfam; PF00483; NTP_transferase; 1.
KW Amino-acid biosynthesis; Translation regulation.
SQ SEQUENCE 452 AA; 50407 MW; 18314D39464233F9 CRC64;
Query Match 12.5%; Score 233.5; DB 1; Length 452;

```

```

Best Local Similarity 21.5%; Prod. No. 4.6e-11;
Matches 96; Conservative 76; Mismatches 144; Indels 137; Gaps 17;

Qy 2  KALIVGCGFCRTLRPLTSLSPKPLVDLPANKPMILHQLQIEALKEGVTEVVLAINTRPEVMI 61
Db 4  QAVYMAVGGSGRTDITSSIFKPLPLFYGNKPLIYPLNLLERYGFEVIVYTT--KDVQK 61
Qy 62  NFKLDPEDKLGITTCQGEPELCTAGPLALARDKLADSCGPPFVLNSDVISEVDPAL 121
Db 62  ALCAEFKMKKLPDILVCIPDEADMGTADSLRHYPKLA-----TDVLVLGCDUITDVALHEV 117
Qy 122 IKFHKHGGEATIMYTRVDE-----PSKYGV-----VYMEATGRVERFV----- 161
Db 118 VDLPRVADSLMLRKQCGSTEDVPQCKKKKTVQEDDFICVDSTSKRLFLFMANEADLD 177
Qy 162 -----EKPKI-FVGNKINAGIYLLNPSVLDRI--ELRPTSIKEVFPQI----- 202
Db 178 EELVIKGSIIQKHPRIHFQGLVDAHLVCLKKYVVDFLMEKNSITSIRSELPILVRKQF 237
Qy 203 --AADQ-----LYAMV-----LPGF-----WMDVGOPR 224
Db 238 SSASSQORQEDKEEDLKKKPKSLDIYSFIKKDNTLTLPYDACWNAFRDRKWDLSRSQ 297
Qy 225 DYITGLRLYLDSTKKSAAKAT-GAHVVG-----VLVHESAKITGECGLI 269
Db 298 -----VRCYVHIKGLCSRVSTGLYMEANRQVPKLLSVLCPEESMIHPSAQIANKHLI 352
Qy 270 GDP-----VAIGPGCVVEDGVRLSRCTVNRGVRIRKKHACISNIIGW 312
Db 353 GADSLGSDTQVGEKSKSRKSVIGSSCVIDRVTNVCNLLMNSVTVGSGSIHGSVFN 412
Qy 313 STV-----GOWARIE-----NMTILGED 330
Db 413 AVVEAGAEIRDCLIGSGRIEAKAKRMNEIVGND 447

RESULT 6
E2BG_HUMAN STANDARD: PRT: 452 AA.
AC Q9NR50; Q9H850;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Translation initiation factor eIF-2B gamma subunit (eIF-2B GDP-GTP
DE exchange factor).
DE EIF2B3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20359747; PubMed=10900014;
RA Krueger M., Beger C., Li Q.-X., Welch P.J., Tritz R., Leavitt M.,
RA Barber J.R., Wong-Staal F.;
RT Identification of eIF2B gamma and eIF2 gamma as cofactors of
RT hepatitis C virus internal ribosome entry site-mediated translation
RT using a functional genomics approach.;
RL PROC. Natl. Acad. Sci. U.S.A. 97:8566-8571(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Ishigai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa M., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wagatsuna M., Hosoliri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuho Y.,
RA Ninomiya K., Iwayanagi T.;
RP NEDO human cDNA sequencing project.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CATALYZES THE EXCHANGE OF EUKARYOTIC INITIATION FACTOR
CC -!- 2-BOUND GDP FOR GTP.
CC -!- SUBUNIT: COMPLEX OF FIVE DIFFERENT SUBUNITS; ALPHA, BETA, GAMMA,

```

```

CC      DELTA AND EPSILON.
CC      -!- SIMILARITY: BELONGS TO THE EIF-2B GAMMA/EPSILON SUBUNITS FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AF257077; AAF91351.1; -.
DR      EMBL; AK024006; BAB14770.1; -.
DR      Genew; HGNC:3259; EIF2B3.
DR      MIM; 606273; -.
DR      InterPro: IPR001825; NTP_transferase.
DR      Pfam; PF00483; NTP_transferase; 1.
KW      Amino-acid biosynthesis; Translation regulation.
FT      CONFLICT 356 356 S -> G (IN REF. 2).
SQ      SEQUENCE 452 AA; 50240 MW; 6F73137F59E52773 CRC64;

      Query Match      12.3%; Score 229.5; DB 1; Length 452;
      Best Local Similarity 20.9%; Pred. No. 9.5e-11;
      Matches 95; Conservative 78; Mismatches 145; Indels 137; Gaps 16;

QY      2 KALILVGGFGTRLRPLTSLFPEKPLVDFAKPMILHQLAEALKEVGVTEVVLAINRPEVMI 61
DB      :|::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
QY      62 NFLKDFEDKLGITITCSQETEPITAGPLALARDKADGSGQFFVLNSDVISEYFPFABL 121
DB      :|::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
QY      62 ALCAEFKMKRPDIVCIPDDADMGTSLSRYTPKLK---TDVLVLSCLDITDVALHEV 117
QY      122 IKFKHCHGGEATIMVK-----VDEPSKYGVVVMVEATGR 156
DB      :|::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
QY      118 VDLFRAYDASLAMLMRKGQDSIEPPVGQKGGKAYEQRDFIGVDSTGRKLLFMANEADILD 177
QY      157 VERVF-----EKPI-FVGKNKINAGIYLLNPSVLDRIELRP--TSIEKEVFPQI----- 202
DB      |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
QY      178 EELVYKSILOKHPRIHFHTGLVDLHLCYKIIYVDFLWENGSIYSIRSELYPIVLRKQF 237
QY      203 --AADQQ-----LYAMVLPG-----FWMVDVGQPR 224
DB      |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
QY      238 SSASSQOQGEKEEDLKKELKSLDIYSFIKEANTLNLAPYDACWACGRDWRDLRSQ 297
QY      225 DYITGLRILYLSIRKKSAAKLAT-GAHVGN-----VLVHESAKI----- 263
DB      |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
QY      298 -----VRCYVHIMKEGLCSRVSTGLYMEANRQVPKLLSALCPPEPPVHSSAQIVSKHLV 352
QY      264 GEGCLIGPDV-----AIGPCVVEDGVRLSRTVMRGVRIKKHACISNLSIGWH 312
DB      |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
QY      353 GVDSLIGPETQIGKSSIKRSYIGSSCLIKDRVTITCLLMSVTVEEGSNIGQSVICNN 412
QY      313 ST-----VGOWARIE-----NWTILGED 330
DB      |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
QY      413 AVIEKGADIKDCLIGSGGRIEAKARVNEIVGND 447
DB      |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|

RESULT 7
ID      SPISL_BACSU
AC      P39629; STANDARD; PRT; 246 AA.
DT      01-FEB-1995 (Rel. 31, Created)
DT      01-FEB-1995 (Rel. 31, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Spore coat polysaccharide biosynthesis protein spsI.
GN      SPSI OR IPA-71D.
OS      Bacillus subtilis.
OC      Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX      NCBI_TaxID=1423;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-168;

```

```

RX      MEDLINE=95020537; PubMed=7934828;
RA      Glaser P., Kunst F., Arnaud M., Coudart M.P., Gonzales W.,
RA      Hulio M.F., Ionescu M., Lubochinsky B., Marcelino L., Moszer I.,
RA      Presecan E., Santana M., Schneider E., Schweizer J., Vertes A.,
RA      Rapoport G., Danchin A.;
RT      "Bacillus subtilis genome project: cloning and sequencing of the 97
RL      kb region from 325 degrees to 333 degrees.";
RL      Mol. Microbiol. 10:371-384(1993).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN-168;
RX      MEDLINE=98044033; PubMed=9384377;
RA      Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA      Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA      Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA      Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA      Chol S.K., Codani J.J., Connerthon I.F., Cummings N.J., Daniel R.A.,
RA      Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA      Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA      Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA      Gim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA      Guisepi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA      Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
RA      Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA      Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA      Kurita K., Lapidus A., Lardinols S., Lauber J., Lazarevic V.,
RA      Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
RA      Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA      Noone D., O'Reilly M., Ogawa K., Ogilwara A., Oudega B., Park S.H.,
RA      Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA      Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA      Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA      Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA      Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA      Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA      Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA      Totsu Y., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA      Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA      Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA      Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT      "The complete genome sequence of the Gram-positive bacterium Bacillus
RT      subtilis.";
RL      Nature 390:249-256(1997).
CC      -!- PATHWAY: Spore coat polysaccharide biosynthesis.
CC      -!- SIMILARITY: BELONGS TO THE GLUCOSE-1-PHOSPHATE
CC      THYMIDYLTRANSFERASE FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; X73124; CAAS1627.1; -.
DR      EMBL; Z59123; CAB15810.1; -.
DR      Subtilist; BG10617; spsI.
DR      InterPro: IPR001825; NTP_transferase.
DR      Pfam; PF00483; NTP_transferase; 1.
KW      Transferase; Kinase; Nucleotidyltransferase; Complete proteome.
SQ      SEQUENCE 246 AA; 27773 MW; 4348B12FF9347D6D CRC64;

      Query Match      12.2%; Score 228; DB 1; Length 246;
      Best Local Similarity 29.2%; Pred. No. 5.9e-11;
      Matches 81; Conservative 44; Mismatches 106; Indels 46; Gaps 10;

QY      1 MKALILVGGFGTRLRPLTSLFPEKPLVDFAKPMILHQLAEALKEVGVTEVVLAINRPEVM 60
DB      ||::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
DB      1 MKGVILAGGNGSLMPLTAVNKHLLPVGPYPMIYWSIMKLQEAQIKILLI--SQEHM 58
QY      61 INFLK--DFEDKLGITITCSQETEPITAGPLALARDKADGSGQFFVLNSDVISE--- 115
DB      | |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|

```

```

Db 59 POFYKLLNGEELGVITYOVQPAASISDGLSYAKRTKES---FILLGLGNIFEDSL 115
QY 116 YPAEILKHKHCHGGEATIMVTKVDESKYGVVVMEEATGRVERFVKPIFVGNKINAG 175
DB 116 KPYTE--RFEQ-QCKGAKVLLKVEDDPERFGIAEIDKKNRIRSIIEKPEOPTNLCVTG 172
QY 176 IYLLNSVLDRIELRPTISIEKEVFPQTAADQOLYAM-----VLPGFWMVVGOPRDYIT 228
DB 173 IYNDAEVFSYIEQISPSKRGEL--EITDNNLYIENSQLTVDYLSGWMVDAG-----T 224

QY 229 GLRLVLDSTIRKSAAKLATCAHVGNVNLVHESAKIGE 265
DB 225 HESLYLAS-----QLVHQALNKGQ 243

RESULT 8
GLMU_ECOLI
ID GLMU_ECOLI STANDARD; PRT; 456 AA.
AC P17114; P76746;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Bifunctional glmu protein [includes: UDP-N-acetylglucosamine
DE pyrophosphorylase (EC 2.7.7.23) (N-acetylglucosamine-1-phosphate
DE uridylyltransferase); Glucosamine-1-phosphate N-acetyltransferase
DE (EC 2.3.1.57)].
GN GLMU OR B3730 OR 25228 OR EGS4672.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TAXID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85121806; PubMed=6395859;
RA Walker J.E., Gay N.J., Saraste M., Eberle A.N.;
RT "DNA sequence around the Escherichia coli unc operon. Completion of
RT the sequence of a 17 kilobase segment containing asuA, oric, unc,
RT glms and phos."
RL Biochem. J. 224:799-815(1984).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=93315143; PubMed=7686882;
RA Burland V.D., Plunkett G. III, Daniels D.L., Blattner F.R.;
RT "DNA sequence and analysis of 136 kilobases of the Escherichia coli
RT genome: organizational symmetry around the origin of replication."
RL Genomics 16:551-561(1993).
RN [3]
RP REVISIONS.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoudis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohemorrhagic Escherichia coli O157:H7."
RL Nature 409:528-533(2001).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;

```

```

RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsuda E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:11-22(2001).
RN [6]
RP IDENTIFICATION.
RX MEDLINE=94012475; PubMed=9407787;
RA Mengin-Lecreulx D., van Heijenoort J.;
RT "Identification of the glmu gene encoding N-acetylglucosamine-1-
RT phosphate uridylyltransferase in Escherichia coli."
RL J. Bacteriol. 175:6150-6157(1993).
RN [7]
RP CHARACTERIZATION.
RX MEDLINE=94364959; PubMed=9083170;
RA Mengin-Lecreulx D., van Heijenoort J.;
RT "Purification of glucosamine-1-phosphate acetyltransferase and
RT N-acetylglucosamine-1-phosphate uridylyltransferase activities of
RT Escherichia coli: Characterization of the glmu gene product as a
RT bifunctional enzyme catalyzing two subsequent steps in the pathway
RT for UDP-N-acetylglucosamine synthesis."
RL J. Bacteriol. 176:5788-5795(1994).
RN [8]
RP CHARACTERIZATION.
RX MEDLINE=96140233; PubMed=8555230;
RA Gehring A.M., Lees W.J., Mindiola D.J., Walsh C.T., Brown E.D.;
RT "Acetyltransferase precedes uridylyltransferase in the formation of
RT UDP-N-acetylglucosamine in separable active sites of the bifunctional
RT Glmu protein of Escherichia coli."
RL Biochemistry 35:579-585(1996).
CC -1- FUNCTION: BIFUNCTIONAL ENZYME RESPONSIBLE FOR THE ACETYLATION OF
CC GLC-N-1-P TO GIVE GLCNAC-1-P AND THE SYNTHESIS OF UDP-GLCNAC.
CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + D-glucosamine 1-phosphate = CoA +
CC N-acetyl-D-glucosamine 1-phosphate.
CC -1- CATALYTIC ACTIVITY: UTP + N-acetyl-alpha-D-glucosamine 1-phosphate
CC - diphosphate + UDP-N-acetyl-D-glucosamine.
CC -1- PATHWAY: PEPTIDOGLYCAN BIOSYNTHESIS; LIPOPOLYSACCHARIDE
CC BIOSYNTHESIS (LIPID A BIOSYNTHESIS).
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE
CC CYSE/LACA/LPXA/NOFL FAMILY OF ACETYLTRANSFERASES. COMPOSED OF
CC MULTIPLE REPEATS OF [LIV]-G-X(4).
CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT THAT CREATES TWO ORFS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X01631; CAA25784.1; -
DR EMBL; L10328; AAA62082.1; ALT_FRAME.
DR EMBL; L10328; AAA62081.1; ALT_FRAME.
DR EMBL; AE000450; AAC70753.1; -
DR EMBL; AE005605; AAG58933.1; -
DR EMBL; AP002566; BAB38095.1; -
DR EMBL; EG11198; glmu.
DR EMBL; IP001451; Hexapep.transf.
DR InterPro; IPR001825; NTP_transferase.
DR Pfam; PF00132; hexapep; 7.
DR Pfam; PF00483; NTP_transferase; 1.
DR TIGRfams; TIGR01173; glmu; 1.
DR PROSITE; PS00101; HEXAPEP_TRANSFERASES; 1.
KW Transferase; Acyltransferase; Nucleotidyltransferase;
KW Multifunctional enzyme; Peptidoglycan synthesis; Cell wall; Repeat;
KW Complete proteome.
FT CONFLICT 186 187
FT SEQUENCE 456 AA; 40190 MW; 2P3C5C84F673C8A3 CRC64;

```

```
Query Match 12.2%; Score 227; DB 1; Length 456;
Best Local Similarity 24.4%; Pred. No. 1.5e-10;
Matches 100; Conservative 55; Mismatches 158; Indels 96; Gaps 16;

Qy 1 MKALLIVGGGTRLRPLTSLFPPKPLVDFAKPMILHQEALKEVGVTEVVLAINYRPEVM 60
Db 6 MSVILAAGKGRMYS---DLPKVLHTLAGKAMVQHVIDAANELGAARVHLVYGHGGDL 62

Qy 61 INFLKDFEDKIGITTCSETPEPLGTAGPLAARDKLADSGSGOPFVLNSDVLSEYPPAE 120
Db 63 KOALKD--DNLNWL-----QAEQLGTGHAMQQA-----APFFADDEDILMLYGDVP 107

Qy 121 LIKFH-----KCHGEATTMTVKDEPSKYGVVME--BATGRVER---FVEKPKIF 167
Db 108 LISVETLQRLDRAKQGG-IGLLTVKLDPTGYGRITRENGKVTGIVEHKDATDEORQI- 165

Qy 168 VGNKINAGIYLLN-----PSVLDRI 187
Db 166 --QEINTGILTANGAMKRWLAKLTNNNAQGEYITDIILALAYQEGREIVAVHPQRLSEV 223

Qy 188 ELRPTSIEKEVFPQTAADQOLYAMVLPFGFWDVGGQPROYIT-GLRLYLDLSIRKKSAAKLA 246
Db 224 EGVNRLQLSLRERYQSEAEKLLAGVYL-----RDPAREDLRGLTHGRD---VEID 275

Qy 247 TGAHVGVNVLVHESAKIGEGC-----LIGPDVAIGPGCVVDEGVRLSRCTV-----MRG 295
Db 276 TNVIEGNTVTLGHRVKGITGCVKNSVIGDDCEISFYTVVEDANLAACTIGPFARLRPG 335

Qy 296 VRIKHACISNIIIGWSTVGQWARIENMTILGEDVHVGVDEYVSGVV 344
Db 336 AELLEGAHVGNFVEMKARLGRKAGHLYLG-DAEIGDNVNIAGTI 383

RESULT 9
E2BE_YEAST STANDARD; PRT; 712 AA.
ID AC P32501.
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Translation initiation factor eIF-2B epsilon subunit (eIF-2B GDP-GTP
DE exchange factor) (Guanine nucleotide exchange factor subunit GCD6)
DE (GCD complex subunit GCD6).
DE GCD6 OR TIF225 OR YDR211W OR YD8142.12 OR YD8142B.03.
GN Saccharomyces cerevisiae (Baker's yeast).
OS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93180841; PubMed=8441423;
RA Bushman J.L., Asuru A.I., Matts R.L., Hinnebusch A.G.;
RT "Evidence that GCD6 and GCD7, translational regulators of GCN4, are
RT subunits of the guanine nucleotide exchange factor for eIF-2 in
RT Saccharomyces cerevisiae."
RL Mol. Cell. Biol. 13:1920-1932(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RA Oliver K., Shore L., Harris D., Barrell B.G., Rajandream M.A.,
RA Walsh S.V.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: SUBUNIT OF THE GUANINE NUCLEOTIDE EXCHANGE FACTOR FOR
CC EIF-2. REQUIRED TO REPRESS GCN4 TRANSLATION UNDER NONSTARVATION
CC CONDITIONS. GCD6 AND GCD7 REPRESS GCN4 EXPRESSION AT THE
CC TRANSLATIONAL LEVEL BY ENSURING THAT RIBOSOMES WHICH HAVE
CC TRANSLATED UORFI WILL REINITIATE AT UORF2, -3, OR -4 AND THUS FAIL
CC TO REACH THE GCN4 START SITE.
CC -1- SUBUNIT: COMPLEX OF FIVE DIFFERENT SUBUNITS; ALPHA (GCN3), BETA
CC (GCD7), GAMMA (GCD1), DELTA (GCD2) AND EPSILON (GCD6).
CC -1- SIMILARITY: BELONGS TO THE EIF-2B GAMMA/EPSILON SUBUNITS FAMILY.
-----
```

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: L07115; AAA65498.1; -
DR EMBL: Z68194; CAA92354.1; -
DR EMBL: Z68195; CAA92362.1; -
DR PIR: S30776; S30776.
DR PIR: A48156; A48156.
DR SGD: S0002619; GCD6.
DR InterPro: IPR001451; Hexapep_transf.
DR InterPro: IPR003307; eIF5C.
DR Pfam: PF00132; hexapep; 4.
DR Pfam: PF02020; W2; 1.
DR SMART: SM00515; eIF5C; 1.
KW Amino-acid biosynthesis; Translation regulation.
SQ SEQUENCE 712 AA; 81161 MW; EFE87F6AE2941619 CRC64;

Query Match 12.0%; Score 224.5; DB 1; Length 712;
Best Local Similarity 22.3%; Pred. No. 4.1e-10;
Matches 92; Conservative 76; Mismatches 155; Indels 89; Gaps 12;

Qy 1 MKALLIVGGGTRLRPLTSLFPPKPLVDFAKPMILHQEALKEVGVTEVVLAINYRPEVM 60
Db 27 LQAVLTDSYETRFMPLTAVKPRCLLPANVPLIETLEFLAKAGVHEVFLICSHANOI 86

Qy 61 INFLKDFEDKIGITTCSETPEPLGTAGPLAARDKLADSGS---QPFVFLNSDV 112
Db 87 NDYIENSKWNLWSPFPKITTIMSPEARCTGD-----VWRD--LDNRGIITGDFILVSGDV 139

Qy 113 ISEYPPFAELIKFHK-----CHGEATIMTVKVDPSKYGVVMEETATGRVER 159
Db 140 LTNIDFSKMLEFHKHKLQDKDHTWCLSKASTYPTRTIEPAF---VLDKSTSRCY 196

Qy 160 FVEKPKIFGVNKNAGIYLLNPSVLDRI----- 188
Db 197 YQDLPLPSSREKTSIQI---DPELLDNVDFVIRNDLIDICDICTSHVPLIFOENFDYQ 253

Qy 189 -LRPTSTEKEVFPQIAADQOLYAMVLPFGFWDVGGQPRODYITGLR-----LYLDSIR 238
Db 254 SLR-TDFVKGVISSDILGKHIIAYLTDEYAVRVESWQTYDTISQDFLGRWCYPLVLDNSI 312

Qy 239 KSAAKLATGAHVY--GNVLVHESAKIGEGCLIGPDVAIGPGCVVDEGVRLSRCTVMRGV 296
Db 313 QDDQTVSYESRHIYKEDVVLQAQCKIGKCTAIGSGTKIGEGTKIENSVIGRNCQIGENI 372

Qy 297 RIK-----KHACISNIIIGWSTVGQWARIENMTILGEDVHVGVDEV 337
Db 373 RIKNSFIWDDCIIGNSIIDSLSIASNATLGSNVRLNDGCIIGFNVKIDNNM 424

RESULT 10
GLGC_BACST STANDARD; PRT; 387 AA.
ID AC O08326;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Glucose-1-phosphate adenylyltransferase (EC 2.7.27) (ADP-glucose
DE synthase) (ADP-glucose pyrophosphorylase).
GN GLGC.
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacilliales; Geobacillus.
OX NCBI_TaxID=1422;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TRBE14;
RX MEDLINE=97386405; PubMed=9244254;
RA Takata H., Takana T., Okada S., Takagi M., Imanaka T.;
```

"Characterization of a gene cluster for glycogen biosynthesis and a heterotetrameric ADP-glucose pyrophosphorylase from *Bacillus stearothermophilus*."  
 RL J. Bacteriol. 179:4689-4698(1997).  
 CC -1- CATALYTIC ACTIVITY: ATP + alpha-D-glucose 1-phosphate -> diphosphate + ADP-glucose.  
 CC -1- PATHWAY: Glycogen biosynthesis; first step.  
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL AND PLANTS GLUCOSE-1-PHOSPHATE ADENYLYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: D87026; BAA19589.1; -  
 DR InterPro: IPR001825; NTP\_transferase.  
 DR Pfam: PF00483; NTP\_transferase; 1.  
 DR PROSITE: PS00808; ADP\_GLC\_PYROPHOSPH\_1; 1.  
 DR PROSITE: PS00809; ADP\_GLC\_PYROPHOSPH\_2; 1.  
 DR PROSITE: PS00810; ADP\_GLC\_PYROPHOSPH\_3; 1.  
 KW Glycogen biosynthesis; Transferase; Nucleotidyltransferase.  
 SQ SEQUENCE 387 AA; 43268 MW; 3660DB962442D97B CRC64;  
 -----  
 Query Match 12.0%; Score 223.5; DB 1; Length 387;  
 Best Local Similarity 23.8%; Pred. No. 2.3e-10;  
 Matches 90; Conservative 64; Mismatches 157; Indels 67; Gaps 14;  
 QY 3 ALLVGGFGTRRLRLTLSPFKPLVDFAKPMIL-HOTEALKEVGTEVVLAINRPEVMI 61  
 DB 7 AMLLAGGSGSKRLSLTNIAPKPAVFGGKYRIIDFTLSNCTNSGIDTVGLVTQYOPILLH 66  
 QY 62 NFLK-----DFEDKL-GITITCSQETP-----LGTAGPLALARDKLDGSGGQFFVLNS 110  
 DB 67 SYIGISAWDLDRNGGVTLPVSYSGVKKVWECTANAVQNTNYTEQVNDPVDVLSG 126  
 QY 111 DVISEYFPFABLRFKHKHCGGEATIMTKV--DEPSKYGVVMEATGRVERFVEKPKIFV 168  
 DB 127 DHIYKMDYQHMLDYHTAKQADVTSIVTEWEEASRFGIMNTNEMIVE-FAEKPAEPK 185  
 QY 169 GNKINAGIYLLNPSVLDRI----ELRPTS---IEKEVFPQTAADQOL-YAMVLPFGMDV 220  
 DB 186 SNLASMGIYIFNPLKQYLQIDNANPHSHDFGKVDYIPMLLREKKRFPAYFPFGYKDV 245  
 QY 221 CQRPD-----YITGLRLYLDSTRKSAKLAATGCAHVGVNVLVHESAKI 263  
 DB 246 GTVKSLWEANMDDLLENDELDRSWRIY--SYNPNOPQYISPEAEVSDSLVNE----- 299  
 QY 264 GEGCLIGPDVAIGFCYVEDGYRLSCTYMRGVRKIKKACISNLIIGHSTVGOMARLEN 323  
 DB 300 -----GCVVEGTV--ERSVLFQGVGRIGKAVKVESVIMPGAAYVSEGAVYER 343  
 QY 324 MTILGEDV-----HVCDE 336  
 DB 344 AIVTPDSIIIPHSYCYPE 361  
 -----  
 RESULT 11  
 RFBA\_RHISN STANDARD; PRT; 286 AA.  
 AC P55464;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE Probable glucose-1-phosphate thymidyltransferase (EC 2.7.7.24)  
 DE (dUDP-glucose synthase) (dUDP-glucose pyrophosphorylase).  
 GN Y4GH.  
 OS *Rhizobium* sp. (strain NGR234).  
 OC Plasmid sym pNGR234a.  
 CC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 RP SEQUENCE FROM N.A.

Rhizobiaceae; *Rhizobium*.  
 NCBI\_TaxID=394;  
 [1]  
 SEQUENCE FROM N.A.  
 RX MEDLINE=97303956; PubMed=9163424;  
 RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A., Perret X.;  
 RA "Molecular basis of symbiosis between *Rhizobium* and legumes."; Nature 387:394-401(1997).  
 RL -1- CATALYTIC ACTIVITY: dTTP + alpha-D-glucose 1-phosphate -> diphosphate + dTDP-glucose.  
 CC -1- PATHWAY: dTDP-L-RHAMNOSE BIOSYNTHESIS  
 CC -1- SIMILARITY: BELONGS TO THE GLUCOSE-1-PHOSPHATE THYMIDYLYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AE000074; AAB91682.1; -  
 DR InterPro: IPR001825; NTP\_transferase.  
 DR Pfam: PF00483; NTP\_transferase; 1.  
 DR TIGRFAMS: TIGR01207; TmlA; 1.  
 KW Hypothetical protein; Plasmid; Transferase; Kinase; Nucleotidyltransferase.  
 SQ SEQUENCE 286 AA; 31205 MW; 2FD7BFF7575914F2 CRC64;  
 -----  
 Query Match 11.3%; Score 210; DB 1; Length 286;  
 Best Local Similarity 25.8%; Pred. No. 1.9e-09;  
 Matches 72; Conservative 54; Mismatches 131; Indels 22; Gaps 9;  
 QY 1 MKALLVGGFGTRRLRLTLSPFKPLVDFAKPMILHQIEALKEVGTEVVLAINRPEVM 60  
 DB 1 MKGIIAGSGSTRLPMTLVMSKQILPVYDKPMIPVPTTLMLAGIREILITSTPHHML 60  
 QY 61 INFLKDFEDKDKITTCSETETPLGTAGPLALARDKLDGSGOP--FFVLNSDVISEYPPA 119  
 DB 61 FQALLGDSGQSGISRYAVQSPNGLAQYVIGADFVA---GSPSCLLDGNDIYFGHLQ 117  
 QY 120 ELIKPHKHCGGEATIMTKVDEPSKYGVVMEAEATGRVERFVEKPKIFVGKINAGIYLL 179  
 DB 118 GILQQAALQOGATIFAYHVDPERYGVVEFDEGNALS-IEEKPAAPKSTWAVTGLIFY 176  
 QY 180 NPSVLD-RIELRPTSIEKEVFPQTAADQOLY-----AMVLPGF-HMDVGGQPRDYTG 229  
 DB 177 DSEVVDTAANLKPSARCEV---EITDVNRIVLERCKLVAVILGRGVAMLDGTP-DSLLE 232  
 QY 230 LRLYLDSTRKSAKLAATGAHV---VGNVLVHESAKIGE 265  
 DB 233 AAEVTVLEKRGKQAKVACPEEVALAMGFISVEEFARIAE 271  
 -----  
 RESULT 12  
 YD34\_METJA STANDARD; PRT; 283 AA.  
 ID YD34\_METJA  
 AC Q58730;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Putative UTP-glucose-1-phosphate uridylyltransferase (EC 2.7.7.9)  
 DE (UDP-glucose pyrophosphorylase) (UDPGP) (Alpha-D-glucosyl-1-phosphate uridylyltransferase) (Uridine diphosphoglucose pyrophosphorylase).  
 GN MJ1334  
 OS *Methanococcus jannaschii*.  
 CC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
 CC Methanocaldococcaceae; Methanocaldococcus.  
 NCBI\_TaxID=2190;  
 [1]  
 RP SEQUENCE FROM N.A.



Query Match 11.28; Score 209; DB 1; Length 456;  
Best Local Similarity 23.58; Pred. No. 4e-09;







GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 27, 2002, 05:01:58 ; Search time 82 Seconds  
(without alignments)  
907.110 Million cell updates/sec

Title: US-09-374-967-2  
Perfect score: 1864  
Sequence: 1 MKALILVGGFGRRLRLTLTSL.....GVVLPHKIKSILKPEIVM 361

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTRMBL\_21.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1766	94.7	361	10 Q941T9	Q941T9 oryza sativ
2	1681	90.2	361	10 O22287	O22287 arabidopsis
3	1680	90.1	361	10 Q9ZTW5	Q9ZTW5 solanum tub
4	1674	89.8	361	10 O9C5B8	O9C5B8 arabidopsis
5	1671	89.6	361	10 Q94IA7	Q94IA7 nicotiana t
6	1587.5	85.2	364	10 Q9M2S0	Q9M2S0 arabidopsis
7	1376	73.8	351	10 Q9M0A3	Q9M0A3 arabidopsis
8	1183.5	63.5	360	4 Q9Y5P6	Q9Y5P6 homo sapien
9	1176.5	63.1	364	3 Q74624	Q74624 trichoderma
10	1175	63.0	363	3 Q74484	Q74484 schizosacch
11	1173	62.9	369	5 Q9VN61	Q9VN61 drosophila
12	1170.5	62.8	364	3 Q9P8N0	Q9P8N0 pichia angu
13	1160	62.2	387	4 Q9H7U3	Q9H7U3 homo sapien
14	1121	60.1	361	3 Q9Y725	Q9Y725 candida gla
15	1099.5	59.0	362	3 Q93827	Q93827 candida alb
16	1045	56.1	373	5 O44970	O44970 caenorhabdi

17	967	51.9	379	5 Q9U1C2	Q9U1C2 leishmania
18	953	51.1	379	5 Q8BLW4	Q8BLW4 leishmania
19	789.5	42.4	345	5 Q8SQX7	Q8SQX7 encephalito
20	594.5	31.9	420	11 Q922H4	Q922H4 mus musculu
21	592.5	31.8	420	4 Q9NWC3	Q9NWC3 homo sapien
22	592.5	31.8	420	4 Q96IJ6	Q96IJ6 homo sapien
23	551.5	29.6	399	4 Q9Y5P5	Q9Y5P5 homo sapien
24	550	29.5	411	10 Q9SSG7	Q9SSG7 arabidopsis
25	550	29.5	415	10 Q9C9P3	Q9C9P3 arabidopsis
26	539	28.9	401	5 Q9N4V3	Q9N4V3 caenorhabdi
27	536	28.8	406	10 Q8W4J5	Q8W4J5 arabidopsis
28	523	28.1	438	5 Q8SXU3	Q8SXU3 drosophila
29	521	28.0	448	5 Q9V7E1	Q9V7E1 drosophila
30	514	27.6	394	5 Q9N4V2	Q9N4V2 caenorhabdi
31	509	27.3	831	16 Q9K2P4	Q9K2P4 streptomyce
32	504	27.0	842	16 Q8YU15	Q8YU15 anabaena sp
33	493.5	26.5	385	10 Q9S3J4	Q9S3J4 arabidopsis
34	492	26.4	831	16 Q66933	Q66933 aquifex aeo
35	490.5	26.3	414	3 Q60064	Q60064 schizosacch
36	489	26.2	392	17 Q8TL99	Q8TL99 methanosarc
37	488	26.2	359	16 P96869	P96869 mycobacteri
38	483	25.9	413	17 Q8U2G7	Q8U2G7 pyrococcus
39	482	25.9	360	16 Q9KZK6	Q9KZK6 streptomyce
40	480	25.8	778	16 Q8R8I4	Q8R8I4 thermoanaer
41	478	25.6	413	17 Q9V037	Q9V037 pyrococcus
42	475	25.5	416	17 Q58775	Q58775 pyrococcus
43	473.5	25.4	361	17 Q980H9	Q980H9 sulfolobus
44	467.5	25.1	357	17 Q96Y83	Q96Y83 sulfolobus
45	467	25.1	349	16 Q8RDG7	Q8RDG7 thermoanaer

ALIGNMENTS

RESULT 1

Q941T9	ID	Q941T9	PRELIMINARY;	PRT;	361 AA.
AC	Q941T9;				
DT	01-DEC-2001	(TRENBLrel. 19, Created)			
DT	01-DEC-2001	(TRENBLrel. 19, Last sequence update)			
DT	01-JUN-2002	(TRENBLrel. 21, Last annotation update)			
DE	Putative GDP-mannose pyrophosphorylase.				
GN	P0005H10.9 OR P0446B05.26.				
OS	Oryza sativa (Rice), and				
OS	Oryza sativa (japonica cultivar-group).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;				
OC	Ehrhartoideae; Oryzeae; Oryza.				
OX	NCBI_TaxID=4530, 39947;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=CV. NIPPONBARE;				
RA	Sasaki T., Matsumoto T., Yamamoto K.;				
RT	"Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC				
RT	clone:P0005H10."				
RL	Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=CV. NIPPONBARE;				
RA	Sasaki T., Matsumoto T., Yamamoto K.;				
RT	"Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC				
RT	clone:P0446B05."				
RL	Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.				
RL	EMBL; AP004127; BAB64272.1; "				
DR	EMBL; AP003251; BAB89577.1; "				
DR	InterPro; IPR001451; Hexapep_transf.				
DR	InterPro; IPR001825; NTP_transferase.				
DR	Pfam; PF00132; hexapep; 4.				
DR	Pfam; PF00483; NTP_transferase; 1.				
DR	PROSITE; PS00101; HEXAPEP_TRANSFERASES; UNKNOWN.1.				
SQ	SEQUENCE 361 AA; 39606 MW; 9E0B29A8107075D2 CRC64;				

Query Match 94.7% Score 1766; DB 10; Length 361;

```
Best Local Similarity 93.6%; Pred. No. 1.4e-136;
Matches 338; Conservative 14; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKALLVGGGCTRLRLPLTSLFPKPLVDFAFKPMILHQIEALKEVGTVTVLAINRYPEVM 60
Db 1 MKALLVGGGCTRLRLPLTSLFPKPLVDFAFKPMILHQIEALKEVGTVTVLAINRYPEVM 60

QY 61 INFLKDFEDKLGITITCSOETBPLGTAGPLALARDKLADGSGQFFVLNSDVISEYFAE 120
Db 61 INFLKDFEDKLGITITCSOETBPLGTAGPLALARDKLADGSGQFFVLNSDVISEYFAE 120

QY 121 LTKFKHCCEATIMVTKVDESKYGVVMEATGRVERFVKPKTFVGNKINAGIYLLN 180
Db 121 LTKFKHCCEATIMVTKVDESKYGVVMEATGRVERFVKPKTFVGNKINAGIYLLN 180

QY 181 PSVLDRLELPTSIKEVEPPOIAADQOLYAMVLPFGWMDVGPDRDYITGLRLYLDLSTKK 240
Db 181 PSVLDRLELPTSIKEVEPPOIAADQOLYAMVLPFGWMDVGPDRDYITGLRLYLDLSTKK 240

QY 241 SAAKLATGAHVGNVIVHESAKICEGCLIGDPAIGCCVVEDGVRSLRCTVMRGVRIKK 300
Db 241 SAAKLATGAHVGNVIVHESAKICEGCLIGDPAIGCCVVEDGVRSLRCTVMRGVRIKK 300

QY 301 HACISNSIIGWSTVGCQWARIENMTILGEDYHVCDEVYSGGVVLPKHEKSKSLKPEIV 360
Db 301 HACISNSIIGWSTVGCQWARIENMTILGEDYHVCDEVYSGGVVLPKHEKSKSLKPEIV 360

QY 361 M 361
Db 361 M 361

RESULT 2
022287 PRELIMINARY; PRT: 361 AA.
ID 022287
AC 022287;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Putative mannose-1-phosphate guanyltransferase
DE (A02G3970/T517.7)
GN T517.7 OR GMPI OR CYTL.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Rounsley S.D., Lin X., Ketchum K.A., Crosby M.L., Brandon R.C.,
RA Sykes S.M., Kaul S., Mason T.M., Kerlavage A.R., Adams M.D.,
RA Somerville C.R., Venter J.C.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Weers B., Thornburg R.W.;
RA "Characterization of the cDNA and gene for the Arabidopsis thaliana
RT GDP-mannose pyrophosphorylase";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lukowitz W., Somerville C.;
RT "Positional cloning of the Arabidopsis cytl gene.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
[4]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Shinn P., Banh J., Bowser L.,
RA Carninci P., Chung M.K., Goldsmith A.D., Hayashizaki Y., Ishida J.,
RA Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M.,
RA Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
Ecker J.R.;
RT "Arabidopsis cDNA clones."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
[5]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
Ecker J.R.;
RT "Arabidopsis cDNA clones."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
[6]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
Ecker J.R.;
RT "Arabidopsis cDNA clones."
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC003000; AAB87126.1; -
DR EMBL; AF076484; AAC78474.1; -
DR EMBL; AF108060; AAD04627.1; -
DR EMBL; AF361812; AAK32825.1; -
DR EMBL; AF428297; AAL16129.1; -
DR EMBL; AY057541; AAL09781.1; -
DR InterPro; IPR001451; Hexaped_transf.
DR InterPro; IPR001825; NTP_transferase.
DR Pfam; PF00132; hexaped; 3.
DR Pfam; PF00483; NTP_transferase; 1.
DR PROSITE; PS00101; HEXAPEP_TRANSFERASES; UNKNOWN_1.
KW Transferase
SQ SEQUENCE 361 AA; 39577 MW; D282B510E22C2F06 CRC64;
Query Match 90.2%; Score 1681; DB 10; Length 361;
Best Local Similarity 88.6%; Pred. No. 1.3e-129;
Matches 320; Conservative 22; Mismatches 19; Indels 0; Gaps 0;

QY 1 MKALLVGGGCTRLRLPLTSLFPKPLVDFAFKPMILHQIEALKEVGTVTVLAINRYPEVM 60
Db 1 MKALLVGGGCTRLRLPLTSLFPKPLVDFAFKPMILHQIEALKEVGTVTVLAINRYPEVM 60

QY 61 INFLKDFEDKLGITITCSOETBPLGTAGPLALARDKLADGSGQFFVLNSDVISEYFAE 120
Db 61 INFLKDFEDKLGITITCSOETBPLGTAGPLALARDKLADGSGQFFVLNSDVISEYFAE 120

QY 121 LTKFKHCCEATIMVTKVDESKYGVVMEATGRVERFVKPKTFVGNKINAGIYLLN 180
Db 121 LTKFKHCCEATIMVTKVDESKYGVVMEATGRVERFVKPKTFVGNKINAGIYLLN 180

QY 181 PSVLDRLELPTSIKEVEPPOIAADQOLYAMVLPFGWMDVGPDRDYITGLRLYLDLSTKK 240
Db 181 PSVLDRLELPTSIKEVEPPOIAADQOLYAMVLPFGWMDVGPDRDYITGLRLYLDLSTKK 240

QY 241 SAAKLATGAHVGNVIVHESAKICEGCLIGDPAIGCCVVEDGVRSLRCTVMRGVRIKK 300
Db 241 SAAKLATGAHVGNVIVHESAKICEGCLIGDPAIGCCVVEDGVRSLRCTVMRGVRIKK 300

QY 301 HACISNSIIGWSTVGCQWARIENMTILGEDYHVCDEVYSGGVVLPKHEKSKSLKPEIV 360
Db 301 HACISNSIIGWSTVGCQWARIENMTILGEDYHVCDEVYSGGVVLPKHEKSKSLKPEIV 360

QY 361 M 361
Db 361 M 361
```

```
Db 361 M 361
RESULT 3
O92TW5 PRELIMINARY; PRT; 361 AA.
AC O92TW5;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE GDP-mannose pyrophosphorylase.
GN GMP.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RA Keller R., Kossmann J., Willmitzer L.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF022716; AAD01737.1; -.
DR InterPro: IPR001451; Hexapep_transf.
DR Pfam: PF00483; NTP_transferase.
DR Pfam: PF00132; hexapep; 3.
DR PROSITE: PS00101; NTP_transferase: 1.
DR PROSITE: PS00101; HEXAPEP_TRANSFERASES; UNKNOWN_1.
SQ SEQUENCE 361 AA; 39578 MW; C12A9C4A43BEE4EC CRC64;

Query Match 90.1%; Score 1680; DB 10; Length 361;
Best Local Similarity 87.8%; Pred. No. 1.6e-129;
Matches 317; Conservative 27; Mismatches 17; Indels 0; Gaps 0;

QY 1 MKALILVGGFGTRLRPLTLSPFKPLVDFANKPMILHQIEALKEVGVTEVVLAINYRPEVM 60
Db 1 MKALILVGGFGTRLRPLTLSPFKPLVDFANKPMILHQIEALKEVGVTEVVLAINYRPEVM 60
QY 61 INFLKDFEDKLGITITCSQETPLGTAGPLALARDKLADGSGOPFFVLNSDVISEYPPAE 120
Db 61 INFLKDFEASLGITITCSQETPLGTAGPLALARDKLADGSGOPFFVLNSDVISEYPPKE 120
QY 121 LIKFHKCHGGEATIMTKVDEPSKYGVVMEETGRVERFEKPKIFVGNKINAGIYLLN 180
Db 121 LIKFHKCHGGEATIMTKVDEPSKYGVVMEETGRVERFEKPKIFVGNKINAGIYLLN 180
QY 181 PSVLDRIELRPTSTIEKEVFPQIAADQOLYAMVLPGFWMVDGQPRDYITGLRLYLDLSIRKK 240
Db 181 PSVLDRIELRPTSTIEKEVFPQIAADQOLYAMVLPGFWMVDGQPRDYITGLRLYLDLSIRKK 240
QY 241 SSKLATGAHVGNVLVHESAKIGEGCLIGDPDVAIGPGCVVEDGVRLSRTVMRGVRIKK 300
Db 241 SSPKLAGSPHIVGNVIVDESAGIKGEGCLIGDPDVAIGPGCVIESGVRLSRTVMRGVRIKK 300
QY 301 HACISNSIIGHSTVGQWARIENMTILGEDVHVCDEVYNSGGVVLPHKEIKSSILKPEIV 360
Db 301 HACISGSIIGHSTVGQWARIENMTILGEDVHVCDEVYNSGGVVLPHKEIKSSILKPEIV 360
QY 361 M 361
Db 361 M 361

RESULT 4
O9C5B8 PRELIMINARY; PRT; 361 AA.
AC O9C5B8;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE GDP-mannose pyrophosphorylase.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
```

```
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Avila C.;
RT "Genes responding to phosphate starvation placed together in
RT Arabidopsis genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ275979; CAC35355.1; -.
DR InterPro: IPR001451; Hexapep_transf.
DR InterPro: IPR001825; NTP_transferase.
DR Pfam: PF00132; hexapep; 3.
DR Pfam: PF00483; NTP_transferase: 1.
DR PROSITE: PS00101; HEXAPEP_TRANSFERASES; UNKNOWN_1.
SQ SEQUENCE 361 AA; 39583 MW; EF36E8FD7D0599D9 CRC64;

Query Match 89.8%; Score 1674; DB 10; Length 361;
Best Local Similarity 88.4%; Pred. No. 4.8e-129;
Matches 319; Conservative 22; Mismatches 20; Indels 0; Gaps 0;

QY 1 MKALILVGGFGTRLRPLTLSPFKPLVDFANKPMILHQIEALKEVGVTEVVLAINYRPEVM 60
Db 1 MKALILVGGFGTRLRPLTLSPFKPLVDFANKPMILHQIEALKEVGVTEVVLAINYRPEVM 60
QY 61 INFLKDFEDKLGITITCSQETPLGTAGPLALARDKLADGSGOPFFVLNSDVISEYPPAE 120
Db 61 INFLKDFETFKLEIKITCSQETPLGTAGPLALARDKLADGSGOPFFVLNSDVISEYPLKE 120
QY 121 LIKFHKCHGGEATIMTKVDEPSKYGVVMEETGRVERFEKPKIFVGNKINAGIYLLN 180
Db 121 ILEFHKSHGGEASIMTKVDEPSKYGVVMEESTGRVERFEKPKLYVGNKINAGIYLLN 180
QY 181 PSVLDRIELRPTSTIEKEVFPQIAADQOLYAMVLPGFWMVDGQPRDYITGLRLYLDLSIRKK 240
Db 181 PSVLDRIELRPTSTIEKEVFPQIAADQOLYAMVLPGFWMVDGQPRDYITGLRLYLDLSIRKK 240
QY 241 SAKLATGAHVGNVLVHESAKIGEGCLIGDPDVAIGPGCVVEDGVRLSRTVMRGVRIKK 300
Db 241 SPAKITSGPHIVGNVLVDEATITGEGCLIGDPDVAIGPGCVIESGVRLSRTVMRGVRIKK 300
QY 301 HACISNSIIGHSTVGQWARIENMTILGEDVHVCDEVYNSGGVVLPHKEIKSSILKPEIV 360
Db 301 HACISGSIIGHSTVGQWARIENMTILGEDVHVCDEVYNSGGVVLPHKEIKSSILKPEIV 360
QY 361 M 361
Db 361 M 361

RESULT 5
O94IA7 PRELIMINARY; PRT; 361 AA.
AC O94IA7;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE GDP-D-mannose pyrophosphorylase (EC 2.7.7.22).
GN GMPASE.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RA Tabata K., Takaoka T., Tateishi Y., Esaka M.;
RT "Gene expression of ascorbic acid-biosynthetic and -oxidizing enzymes
RT and ascorbic acid content in tobacco.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB066279; BAB62108.1; -.
DR InterPro: IPR001451; Hexapep_transf.
DR InterPro: IPR001825; NTP_transferase.
DR Pfam: PF00132; hexapep; 3.
```

DR Pfam: PF00483; NTP\_transferase; 1.  
DR PROSITE: PS00101; HEXAPEP\_TRANSFERASES; UNKNOWN\_1.  
KW nucleotidyltransferase; transferase.  
SQ SEQUENCE 361 AA; 39572 MW; 40ED436D30B99528 CRC64;

Query Match 89.6%; Score 1671; DB 10; Length 361;  
Best Local Similarity 87.3%; Pred. No. 8 5e-129;  
Matches 315; Conservative 28; Mismatches 18; Indels 0; Gaps 0;

OY 1 MKALILVGGFGTRLRPLTSLSPKPLVDFANKPMILHQLAEALKEVGYTEVYLAINYPEYM 60  
DB 1 MKALILVGGFGTRLRPLTSLSPKPLVDFANKPMILHQLAEALKEVGYTEVYLAINYPEYM 60  
OY 61 INFLKDFEDKLGITITCSQETEPGLTAGLADKGLADGSGOPFFVLNSDVISSEYPPAE 120  
DB 61 INFLKDFEDKLGITITCSQETEPGLTAGLADKGLADGSGOPFFVLNSDVISSEYPPAE 120  
OY 121 LKPKHCKHGGEATIMVTKVDPSKYGVVMEATGRVERFVKPKLFGVGNKINAGIYVLN 180  
DB 121 LKPKHCKHGGEATIMVTKVDPSKYGVVMEATGRVERFVKPKLFGVGNKINAGIYVLN 180  
OY 181 PSVLDRIELRPTSLIEKEVFPQIAADQOLYAMVLPFGFMDVGGOPRDYITGLRLYLDLSIRKK 240  
DB 181 PSVLDRIELRPTSLIEKEVFPQIAADQOLYAMVLPFGFMDVGGOPRDYITGLRLYLDLSIRKK 240  
OY 241 SSKPLASGSHVGNVYDESASIKGEGCLIGPDVAIGPGCVWEDGVRLSRTCVMRGVRIKK 300  
DB 241 SSKPLASGSHVGNVYDESASIKGEGCLIGPDVAIGPGCVWEDGVRLSRTCVMRGVRIKK 300  
OY 301 HACISNSIIGHSTVGQWARIENMTILGEDVHVCDVGVNCGVVLPHKKEIKSSILKPEIV 360  
DB 301 HACISNSIIGHSTVGQWARIENMTILGEDVHVCDVGVNCGVVLPHKKEIKSSILKPEIV 360

361 M 361

361 M 361

RESULT 6  
OY Q9M2S0 PRELIMINARY; PRT; 364 AA.  
AC Q9M2S0;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Mannose-1-phosphate guanylyltransferase-like protein.  
GN T22E16.250.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
OX NCBI\_taxid=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Benes V., Wurmbach E., Drzonek H., Ansoerge W., Mewes H.W., Lemcke K.,  
RA Mayer K.F.X., Quetier F., Salanoubat M.,  
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL132975; CAB75917.1; -.  
DR InterPro; IPR001451; Hexapep\_transf.  
DR InterPro; IPR001825; NTP\_transferase.  
DR Pfam; PF00132; hexapep; 3.  
DR Pfam; PF00483; NTP\_transferase; 1.  
DR PROSITE; PS00101; HEXAPEP\_TRANSFERASES; UNKNOWN\_1.  
KW Nucleotidyltransferase; transferase.  
SQ SEQUENCE 364 AA; 39808 MW; 8C2BDAC282CF915D CRC64;

Query Match 85.2%; Score 1597.5; DB 10; Length 364;  
Best Local Similarity 83.2%; Pred. No. 6e-122;  
Matches 303; Conservative 29; Mismatches 26; Indels 3; Gaps 1;

361 M 361

OY 1 MKALILVGGFGTRLRPLTSLSPKPLVDFANKPMILHQLAEALKEVGYTEVYLAINYPE-- 58  
DB 1 MKALILVGGFGTRLRPLTSLSPKPLVDFANKPMILHQLAEALKEVGYTEVYLAINYPEQL 60  
OY 59 -VMINFLKDFEDKLGITITCSQETEPGLTAGLADKGLADGSGOPFFVLNSDVISSEY 117  
DB 61 LVMSKSNVDVETILGKITCSQETEPGLTAGLADKGLADGSGOPFFVLNSDVISDYP 120  
OY 119 FAEILKFKHGGEATIMVTKVDPSKYGVVMEATGRVERFVKPKLFGVGNKINAGIY 177  
DB 121 LEEMIAFHNAHGGEASIMVTKVDPSKYGVVMEATGRVERFVKPKLFGVGNKINAGIY 180  
OY 178 LLNPSVLDRIELRPTSLIEKEVFPQIAADQOLYAMVLPFGFMDVGGOPRDYITGLRLYLDLSI 237  
DB 181 LLNPSVLDRIELRPTSLIEKEVFPQIAAEKLYAMVLPFGFMDIGOPRDYITGLRLYLDLSI 240  
OY 238 RKSAKALATGAHVGNVLVHESAKIGEGCLIGPDVAIGPGCVWEDGVRLSRTCVMRGVR 297  
DB 241 RKSPSKLATGPHILGNVLYDETAIEGEGCLIGFNVAIGFQCVYVESGVRLHNKIKSDILKP 300  
OY 298 IKKHACISNSIIGHSTVGQWARIENMTILGEDVHVCDVGVNCGVVLPHKKEIKSSILKP 357  
DB 301 VKRVACISSIIGHSTVGQWARIENMTILGKNVYVCDIYCNCGVVLHNKIKSDILKP 360  
OY 358 EIVM 361  
DB 361 DIVM 364

RESULT 7

OY Q9M0A3 PRELIMINARY; PRT; 351 AA.  
AC Q9M0A3;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE GDP-mannose pyrophosphorylase like protein (Fragment).  
GN AT4G30570.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
OX NCBI\_taxid=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Lamar B., Stoneking T., Stumpf J., Mewes H.W., Lemcke K.,  
RA Mayer K.F.X.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL161577; CAB79775.1; -.  
DR InterPro; IPR001451; Hexapep\_transf.  
DR InterPro; IPR001825; NTP\_transferase.  
DR Pfam; PF00132; hexapep; 3.  
DR Pfam; PF00483; NTP\_transferase; 1.  
DR NON\_TER 1  
FT SEQUENCE 351 AA; 38681 MW; 88B34E790F5CD3BA CRC64;

Query Match 73.8%; Score 1376; DB 10; Length 351;  
Best Local Similarity 76.3%; Pred. No. 1.2e-104;  
Matches 264; Conservative 44; Mismatches 36; Indels 2; Gaps 2;

OY 1 MKALILVGGFGTRLRPLTSLSPKPLVDFANKPMILHQLAEALKEVGYTEVYLAINYPEY 59  
DB 1 LKALILVGGFGTRLRPLTSLSPKPLVDFGNKPMILHQLAEALKEVGYTEVVLAINHQDEV 60  
OY 60 MINFLKDFEDKLGITITCSQETEPGLTAGLADKGLADGSGOPFFVLNSDVISSEYPPFA 119  
DB 61 MLNFVAIEKKLEKITFTSQTETPLGTAGLADKGLADGSGOPFFVLNSDVISCEYPLL 120

```
QY 120 ELIKFKHCHGEATIMTKVDPSKYGWVVMEEATGRVERFEKPKIFVGNKINAGIYLL 179
Db 121 EMIEFHKTRAEASIMTEVDDPSKYGVVTEEGTARVESFEKPKHFVGNKINAGIYLL 180
QY 180 NPSVLDRLELPTSIIEKEVFPQIAADQOOLYAMVLPFGFMDVGQPRDYITGLRLYLDLSIRK 239
Db 181 SPSVLDRLELRTSIEKEIFKIASEKKLYAMVLPFGFMDVGQPKDYITGQRMVLSNLSRE 240
QY 240 KSAKLATGAHVGNVNLVHESAKIGEGCLIGPDVAIGPGCVVDEGVRLSCTVMRGVRIK 299
Db 241 KTPDELATGDNIGNVNLVHESAVIGEGCLIGPDVVGPGCVIDSGLVRLFGCTVMRGVWIK 300
QY 300 KHACISNIIHGWSTVGQWARIENNTILGEDVHVCD-EVYSGGGV 344
Db 301 EHACISNIVGWDSTVGWRARVENITVLGKDVNVADEVYNSGVYI 346

RESULT 8
QY5P6 PRELIMINARY; PRT; 360 AA.
AC QY5P6;
DT 01-NOV-1999 (TremBrel. 12, Created)
DT 01-NOV-1999 (TremBrel. 12, Last sequence update)
DT 01-DEC-2001 (TremBrel. 19, Last annotation update)
DE GDP-mannose pyrophosphorylase B.
GN GMPB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Matthijs G., Schollen E., Dierickx D.;
RT "human homolog of GDP-mannose pyrophosphorylase.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA Strausberg R.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RC TISSUE-EYE;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF135421; AAD38516.1; -.
DR EMBL; BC001141; AAH01141.1; -.
DR EMBL; BC008033; AAH08033.1; -.
DR InterPro; IPR001451; Hexapep.transf.
DR InterPro; IPR001825; NTP_transferase.
DR Pfam; PF001132; hexapep; 3.
DR Pfam; PF00483; NTP_transferase; 1.
DR PROSITE; PS00101; HEXAPEP_TRANSFERASES; UNKNOWN_1.
SQ SEQUENCE 360 AA; 39862 MW; D359AB87459A6B2B CRC64;

Query Match 63.58; Score 1183.5; DB 4; Length 360;
Best Local Similarity 61.24; Pred. No. 7.6e-89;
Matches 221; Conservative 65; Mismatches 74; Indels 1; Gaps 1;

QY 1 MKALILVGGFGTGLRPLTLSPKPLVDFANKPMILHQIEALKEVGTVEVLAINYRPEVM 60
Db 1 MKALILVGGFGTGLRPLTLSPKPLVDFCNKPLILHQVEALAAAGVDHVLAVSTMSQVL 60
QY 61 INFLKDFEKLGITTCQETPLGTAGLALARDKLADGSGQPPFVLNSDVISEYFAE 120
Db 61 EKEMAKOEORLIGIRISMSHEEPLGTAGLALARDLISE-TADPFVLNSDVICDFPQA 119
QY 121 LIFKFKHCHGEATIMTKVDPSKYGWVVMEEATGRVERFEKPKIFVGNKINAGIYLL 180
Db 120 MVQFRRHHGQEGSLTVTKVEEPSKYGWVVMCEADTGRIHRFEKPKQVFVSNKINAGMITYLS 179
QY 181 PSVLDRLELPTSIIEKEVFPQIAADQOOLYAMVLPFGFMDVGQPRDYITGLRLYLDLSIRK 240
Db 181 PSVLDRLELPTSIIEKEVFPQIAADQOOLYAMVLPFGFMDVGQPRDYITGLRLYLDLSIRK 240
```

```
Db 180 PAVLRRIQLOPTSIIEKEVFPIMAKEGQLYAMELQGFMMDIGQPKDFTGMCLFLQSLRQK 239
QY 241 SAAKLATGAHVGNVNLVHESAKIGEGCLIGPDVAIGPGCVVDEGVRLSCTVMRGVRIK 300
Db 240 QPERICSGGFIGVGNLVDFYSARIGQNCISGPNVSLGPGVVDGVCITRCTVLRDARIS 299
QY 301 HACISNIIHGWSTVGQWARIENNTILGEDVHVCDVYSGVGVLPHPKESILKPEIV 360
Db 300 HSWLESCIVGWRGVRVQWVRMENVTLGEDIYNDVLYLNGASVLPKHSIGESVPEPRII 359
QY 361 M 361
Db 360 M 360

RESULT 9
O74624 PRELIMINARY; PRT; 364 AA.
AC O74624;
DT 01-NOV-1998 (TremBrel. 08, Created)
DT 01-NOV-1998 (TremBrel. 08, Last sequence update)
DT 01-JUN-2001 (TremBrel. 17, Last annotation update)
DE Mannose-1-phosphate guanylttransferase (EC 2.7.13) (MPG1 transferase)
DE (ATP-mannose-1-phosphate guanylttransferase).
GN MPG1.
OS Trichoderma reesei (Hypocrea jecorina).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreales; Hypocreaceae; Hypocrea.
OX NCBI_TaxID=51453;
RN [1]
RP SEQUENCE FROM N.A., AND FUNCTION.
RC STRAIN-ATCC 56765 / QM9414 / RUT C-30;
RX MEDLINE=98309839; PubMed=9644208;
RA Kruszevska J.S., Saloheimo M., Penttila M., Palamarczyk G.;
RT "Isolation of a Trichoderma reesei cDNA encoding GTP: a-D-mannose-1-
RT phosphate guanylttransferase involved in early steps of protein
RT glycosylation.";
RL Curr. Genet. 33:445-450(1998).
CC -!- FUNCTION: ESSENTIAL PROTEIN WHICH MAY BE INVOLVED IN THE
CC REGULATION OF CELL CYCLE PROGRESSION.
CC -!- CATALYTIC ACTIVITY: GTP + ALPHA-D-MANNOSE 1-PHOSPHATE =
CC DIPHOSPHATE + GDP-MANNOSE.
CC -!- SIMILARITY: TO S.TYPHIMURIUM CDP-GLUCOSE PYROPHOSPHORYLASE (REBF).
DR EMBL; U89991; AAC39498.1; -.
DR InterPro; IPR001451; Hexapep.transf.
DR InterPro; IPR001825; NTP_transferase.
DR Pfam; PF00132; hexapep; 4.
DR Pfam; PF00483; NTP_transferase; 1.
DR PROSITE; PS00101; HEXAPEP_TRANSFERASES; UNKNOWN_1.
KW Transferase; Nucleotidyltransferase; Kinase; Cell cycle.
SQ SEQUENCE 364 AA; 40284 MW; 66203BAF15284EBA CRC64;

Query Match 63.11%; Score 1176.5; DB 3; Length 364;
Best Local Similarity 61.4%; Pred. No. 2.9e-88;
Matches 224; Conservative 61; Mismatches 75; Indels 5; Gaps 4;

QY 1 MKALILVGGFGTGLRPLTLSPKPLVDFANKPMILHQIEALKEVGTVEVLAINYRPEVM 60
Db 1 MKGILVGGFGTGLRPLTLSPKPLVEFCNKPIMVHQIEALVAAGVTDIVLAVNYRPEIM 60
QY 61 INFLKDFEKLGITTCQETPLGTAGLALARDKLADGSGQPPFVLNSDVISEYFAE 120
Db 61 EKFLAEYEEKYNINIEFSVESEPLDTAGPLKLA-ERILGKDDSPFFVLNSDVICDYPKE 119
QY 121 LIFKFKHCHGEATIMTKVDPSKYGWVVMEEATGRVERFEKPKIFVGNKINAGIYLL 179
Db 120 LIEFHKAHGEDEGTIVTKVEEPSKYGWVVMKPNHPSRIDRFVEKPFVFNRRINAGMIIF 179
QY 180 NPSVLDRLELPTSIIEKEVFPQIAADQOOLYAMVLPFGFMDVGQPRDYITGLRLYLDLSIRK 239
Db 180 NPSVLKRIELRPTSIIEKETFPAMVADNQLHSPDLEGFMDVGQPKDFLSGTCLYLSLTK 239
QY 240 KSAAKLA--TGAHV-VGNVLVHESAKIGEGCLIGPDVAIGPGCVVDEGVRLSCTVMRGV 296
```

```

DB 240 KGSKELTPTEPIYVGGNNVMIHPSAKIGANCRIQPNVTIGPDVYVGGVQLQRCYLLKGS 299
QY 297 RKXKHCACINSIIGHSTVGQWARIENMTILGDEHVCDDEVSYNGGVLPKHKIKSILK 356
DB 300 KVKDHAWKSTVIGWNSTVGRWARLENVIVLGDVITGDEIYVGGSVLPKHKIKANVDV 359
QY 357 PEIYM 361
DB 360 PAIIM 364

RESULT 10
O74404
ID O74484 PRELIMINARY; PRT; 363 AA.
AC O74484; P78779;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DE Probable mannose-1-phosphate guanylyltransferase (EC 2.7.7.13) (ATP-
DE mannose-1-phosphate guanylyltransferase).
GN SPEC1906.01.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Lyne M., Rajandream M.A., Barrell B.G., Murphy L., Harris D.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=PR745;
RA MEDLINE=98162722; PubMed=9501991;
RA Yoshioka S., Kato K., Nakai K., Okayama H., Nojima H.;
RT "Identification of open reading frames in Schizosaccharomyces pombe
RT CDNA's";
RL DNA Res. 4:363-369(1997).
CC -!- FUNCTION: ESSENTIAL PROTEIN WHICH MAY BE INVOLVED IN THE
CC REGULATION OF CELL CYCLE PROGRESSION (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: GTP + ALPHA-D-MANNOSE 1-PHOSPHATE =
CC DIPHOSPHATE + GDP-MANNOSE.
CC -!- SIMILARITY: TO S.TYPHIMURIUM CDP-GLUCOSE PYROPHOSPHORYLASE (REFB).
DR EMBL: AL031538; CAA20770.1; -.
DR EMBL: D89128; BAA13790.1; -.
DR InterPro: IPR001451; Hexapep_transf.
DR InterPro: IPR001825; NTP_transferase.
DR Pfam: PF00132; hexapep; 4.
DR PROSITE: PS00101; HEXAPEP_TRANSFERASES; UNKNOWN_1.
KW Transferase; Nucleotidyltransferase; Kinase; Cell cycle.
FT CONFLICT 175 176 GI -> VL (IN REF. 2).
SQ SEQUENCE 363 AA: 39719 NW; 55P55FF0ED017C02 CRC64;

Query Match 63.0%; Score 1175; DB 3; Length 363;
Best Local Similarity 62.1%; Pred. NO. 3.8e-68;
Matches 226; Conservative 54; Mismatches 80; Indels 4; Gaps 4;

OY 1 MKALIVGGFGRLRLPTLSFKPKPLVDFANKPMLHIOIALKEVGYTEVVLAINRYPEVM 60
DB 1 MKALIVGGFGRLRLPTLTLLFKPLVDFGNKPKMLHIOIALKEVGYTEVVLAINRYPEIM 60
QY 61 INFLKDFEDKLGITITCSQETPLCPAGPLALARDKLDGSGOFFVLNSDVISYPPFAE 120
DB 61 VEALKKYEKYNVNITFSVENEPLCTAGPLALARDILAK-DHSPFFVLNSDVICEYPPAD 119
QY 121 LKFKHCHGGEATIMYTKVDEFSKYGVV-MEATGRVERFVEKPKIFVGNKKNAGIYLL 179
DB 120 LAAFHKAHCAEGTIVTKVEEPSKYGVVHYVHNSESLIERFVEKPKVFNRSNRRNGGIIV 179
QY 180 NPSVLDRLELRPTSLEKEVFPQIAADDOOLYAMVLPGFWMVDVGPRDIYITGLRLYLSIRK 239

```

```

DB 180 NPSVLDRLELRPTSLEKEVFPQIAADDOOLYAMVLPGFWMVDVGPRDIYITGLRLYLSIRK 239
QY 240 KSAAKLA-TGAHVVGNVLVHESAKIGECCLIGPDVAIGPGCVVDEGVRLSRCTYMGVRI 298
DB 240 HKPEILAPASSNIIGNVLDIPSATIGNKCKIGPNVIGPNVTIGDVRQLQRCAILKSSRV 299
QY 299 KKHACINSIIGHSTVGQWARIENMTILGDEHVCDDEVSYNGGVLPKHKIKSILKP- 357
DB 300 RDHAWKSSITVGWNSTLGSWSRLNVSVLGDVYVVDVDEIYVGGSVLPKHKISANIEYVP 359
QY 358 EIYM 361
DB 360 TIYM 363

RESULT 11
Q9YN61
ID Q9YN61 PRELIMINARY; PRT; 369 AA.
AC Q9YN61;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Cg1123 protein (H102803P) (R549494P).
GN Cg1123.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrial J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottlie P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cheriy J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush P., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spadling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassenaar D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2195-2195(2000).
RN [2]

```





```
DR InterPro; IP001825; NTP_transferase.
DR Pfam; PF00132; hexapep; 3.
DR Pfam; PF00483; NTP_transferase; 1.
DR PROSITE; PS00101; HEXAPEP-TRANSFERASES; UNKNOWN.1.
SQ SEQUENCE 387 AA; 42649 MW; CA3BFAD1848F1FE1 CRC64;

Query Match
Best Local Similarity 62.2%; Score 1160; DB 4; Length 387;
Matches 221; Conservative 65; Mismatches 74; Indels 28; Gaps 2;

QY 1 MKALILVGGFGTGLRLPTLSFPKPLVDFAFKPMILHQIEALKEVGVTEVVLAINRPEVM 60
DB 1 MKALILVGGFGTGLRLPTLSFPKPLVDFAFKPMILHQIEALKEVGVTEVVLAINRPEVM 60

QY 61 INFLKDPEDKLGITTCQETPEPLGTAGPLALARDKLADGGSGOFFVYLNDSVISEYFPAE 120
DB 61 INFLKDPEDKLGITTCQETPEPLGTAGPLALARDKLADGGSGOFFVYLNDSVISEYFPAE 120

QY 121 LKFKHCHGCGEATIVTKVDEPSKYGVVMEATGRVERFVEKPKIFVGNNKINAGIYLLN 180
DB 121 LKFKHCHGCGEATIVTKVDEPSKYGVVMEATGRVERFVEKPKIFVGNNKINAGIYLLN 180

QY 120 MVQFRRHGGEGSILVTKVEEPSKYGVVCEADTGRTHREVEKPVQFVSNKINAGMILS 179
DB 120 MVQFRRHGGEGSILVTKVEEPSKYGVVCEADTGRTHREVEKPVQFVSNKINAGMILS 179

QY 181 PSYLDRIELRPTSIKEVEFQIAADQQLYANVLGCFWMDVQOPRDYITGLRLYLDSTRK 240
DB 181 PSYLDRIELRPTSIKEVEFQIAADQQLYANVLGCFWMDVQOPRDYITGLRLYLDSTRK 240

QY 180 PAVLRIRIQOPTSIKEVEFQIAADQQLYANVLGCFWMDVQOPRDYITGLRLYLDSTRK 239
DB 180 PAVLRIRIQOPTSIKEVEFQIAADQQLYANVLGCFWMDVQOPRDYITGLRLYLDSTRK 239

QY 241 SAAKLATGATCAHVGNVLVHESAKIGEGCLIGPDVAIGCGVYEDGVRSLRCTVMRGVRIK 300
DB 241 SAAKLATGATCAHVGNVLVHESAKIGEGCLIGPDVAIGCGVYEDGVRSLRCTVMRGVRIK 300

QY 240 QPERLGSFGFVGNVLYDPESARIGQNCISGPNVSLGFGVYEDGVCIRCTVLRDAIRS 299
DB 240 QPERLGSFGFVGNVLYDPESARIGQNCISGPNVSLGFGVYEDGVCIRCTVLRDAIRS 299

QY 301 HACISNIIHSHSTVCOWARIENMTILGEDVHVCDEYVSNGGVVLPHKEIKSSILKPEI 359
DB 301 HACISNIIHSHSTVCOWARIENMTILGEDVHVCDEYVSNGGVVLPHKEIKSSILKPEI 359

QY 300 HSWLESCIVGWRKRVQWVSLWAGLGGEGGACACLPDKAYFLLEVRMENVTVLGEDVIV 359
DB 300 HSWLESCIVGWRKRVQWVSLWAGLGGEGGACACLPDKAYFLLEVRMENVTVLGEDVIV 359

QY 334 CDEYVSNGGVVLPHKEIKSSILKPEI 361
DB 334 CDEYVSNGGVVLPHKEIKSSILKPEI 361

QY 360 NDELYLNGASVLPKHSIGESVPEPIIM 387
DB 360 NDELYLNGASVLPKHSIGESVPEPIIM 387

RESULT 14
QY725
ID QY725 PRELIMINARY; PRT; 361 AA.
AC QY725;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE GDP-mannose pyrophosphorylase.
GN VIG9.
OS Candida glabrata (Yeast) (Torulopsis glabrata).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5478;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC2001;
RA Ohta A., Sudoh M.;
RT "Candida glabrata VIG9 gene for GDP-mannose pyrophosphorylase.";
DR EMBL; AB020576; BAA77382.1; -.
DR InterPro; IP001451; Hexapep-transf.
DR InterPro; IP001825; NTP_transferase.
DR Pfam; PF00132; hexapep; 4.
DR Pfam; PF00483; NTP_transferase; 1.
DR PROSITE; PS00101; HEXAPEP-TRANSFERASES; UNKNOWN.2.
SQ SEQUENCE 361 AA; 39298 MW; E255084D2EDB2A1D CRC64;

Query Match
Best Local Similarity 60.1%; Score 1121; DB 3; Length 361;
Matches 211; Conservative 62; Mismatches 87; Indels 2; Gaps 2;

QY 1 MKALILVGGFGTGLRLPTLSFPKPLVDFAFKPMILHQIEALKEVGVTEVVLAINRPEVM 60
DB 1 MKALILVGGFGTGLRLPTLSFPKPLVDFAFKPMILHQIEALKEVGVTEVVLAINRPEVM 60
```

```
QY 61 INFLKDPEDKLGITTCQETPEPLGTAGPLALARDKLADGGSGOFFVYLNDSVISEYFPAE 120
DB 61 VETLQKYKEKGYSTFVETPEPLGTAGPLALA-BKVLKDKNSPFVYLNDSVICYPFKE 119

QY 121 LKFKHCHGCGEATIVTKVDEPSKYGVVMEATGRVERFVEKPKIFVGNNKINAGIYLL 179
DB 121 LADPHKAGGAGTIVATKVDPSKYGVVHDIATFNLDTRFVEKPKFVGVRINAGLYIL 179

QY 180 NPSYLDRIELRPTSIKEVEFQIAADQQLYANVLGCFWMDVQOPRDYITGLRLYLDSTRK 239
DB 180 NPEVDLTLEMTPTSETPTDILVQKSLYSPDLGFWMDVQOPKDFLSGTVLYLNSVSK 239

QY 240 KSAKLATGATCAHVGNVLVHESAKIGEGCLIGPDVAIGCGVYEDGVRSLRCTVMRGVRIK 299
DB 240 KNPEKLTKGDNIVGNVMDPSAKIAASAKVGDVVGIVGNVTIGCVRTSRVVLSDSSIQ 299

QY 300 KHACISNIIHSHSTVCOWARIENMTILGEDVHVCDEYVSNGGVVLPHKEIKSSILKPEI 359
DB 300 DHSLVKSTIVGKSTVGVKWCRLGVTILGDNVYVNGKYLPHKHSISANYPSEAI 359

QY 360 VM 361
DB 360 IM 361

RESULT 15
QY3827
ID QY3827 PRELIMINARY; PRT; 362 AA.
AC QY3827;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE Mannose-1-phosphate guanylyltransferase (EC 2.7.7.13) (ATP-mannose-1-phosphate guanylyltransferase) (CASREL).
GN VIG9 OR SRB1 OR PSAL.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1161;
RX MEDLINE=38455809; PubMed=3782409;
RA Marit S., Walmsley R.M., Staveva L.I.;
RT "Cloning and sequencing of the Candida albicans homologue of SRB1/PSAL/VIG9, the essential gene encoding GDP-mannose pyrophosphorylase in Saccharomyces cerevisiae.";
RN Microbiology 144:2417-2426(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=IFO1060;
RA Ohta A., Sudoh M.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ESSENTIAL PROTEIN WHICH MAY BE INVOLVED IN THE REGULATION OF CELL CYCLE PROGRESSION.
CC -1- CATALYTIC ACTIVITY: GTP + ALPHA-D-MANNOSE 1-PHOSPHATE -> DIPHOSPHATE + GDP-MANNOSE.
CC -1- SIMILARITY: TO S.TYPHIMURIUM CDP-GLUCOSE PYROPHOSPHORYLASE (RPF).
DR EMBL; AF030299; AAC64911.1; -.
DR EMBL; AF030300; AAC64912.1; -.
DR EMBL; AB020596; BAA34807.1; -.
DR InterPro; IP001451; Hexapep-transf.
DR InterPro; IP001825; NTP_transferase.
DR Pfam; PF00132; hexapep; 4.
DR Pfam; PF00483; NTP_transferase; 1.
KW Transferase; Nucleotidyltransferase; Kinase; Cell cycle.
FT VARIANT 197 198 ET -> DP (IN STRAIN 1161).
FT CONFLICT 289 289 Q -> R (IN AAC64912).
FT VARIANT 333 333 E -> Q (IN STRAIN 1161).
SQ SEQUENCE 362 AA; 39975 MW; 052DCC32D0E16923 CRC64;

Query Match
Score 59.08; Score 1099.5; DB 3; Length 362;
```



Best Local Similarity 57.9%; Pred. No. 5.9e-82;			
Matches	210; Conservative	60; Mismatches	90; Indels 3; Gaps 3;
QY	1	MKALILVGGFGTRLRPLTLSPKPLVDFAFKPMILHQIEALKEVGVTEVWLAINYRPEVM	60
Db	1	MKGLILVGGYGTRLRPLTLPLKPLVEFGNRPMLHQIEALAAAGVTDIVLAVNRPVM	60
QY	61	INFLKDFEDKLGITITCSQETPLGTAGPLALARDKLDGSGQPFFVLNSDVISEYFAE	120
Db	61	VSTLKKEEYEGVSGITFSVEEPLGTAGPLKLAEEVLKK-DDSPFFVLNSDVICDYPKE	119
QY	121	LIRFKHCHGGEATIMVTKVDEPSKYGVVWMEAT-GRVEREVEKPKIFVGKNKINAGIYLL	179
Db	120	LADFHKAHGAAGTIVATKVDEPSKYGVIVHDDTPNLIDREVEKPVVEVGNRINAGLYIL	179
QY	180	NPSVLDRIELRPTSIKREVPPIAADQOLYAMVLPGFWMMDVGPDRDYITGLRLYLDLSRK	239
Db	180	NPSVIDLIEMRPTSIKETFPILVQEQLYSFDEGYWMDVGPQKDFLUSGTCCLYLTLSLK	239
QY	240	KSAAKLATGAHV-VGNVLVHESAKIGEGCLIGPDVAIGPGCVVEDGVRLSCTVMRGVRI	298
Db	240	KHPEKLCCKEYVHGVNLIIDPTAKIHPSALIGPNVTIGPNVVGGEGARIQSVLLANSQV	299
QY	299	KKHACISNSTIGWHSTVGQWARIENMTILGEDVHVCDEVYSGNGVVLPHKSIKPE	358
Db	300	KDHAWKSTIVGWNSTRIGKWARTEGTVLGDDVEVKNEIYYNGAKVLPKHSISSNVEKES	359
QY	359	IVM	361
Db	360	IIM	362

Search completed: November 27, 2002, 05:09:35  
Job time : 85 secs